(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 January 2001 (04.01.2001)

PCT

(10) International Publication Number
WO 01/00869 A2

(51) International Patent Classification[7]: C12Q 1/68, 1/68

(21) International Application Number: PCT/IL00/00370

(22) International Filing Date: 23 June 2000 (23.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
130606 23 June 1999 (23.06.1999) IL

(71) Applicant *(for all designated States except US)*: **INTELLIGENE LTD.** [IL/IL]; P.O. Box 35241, 91351 Jerusalem (IL).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **NATHAN, Asher** [IL/IL]; P.O. Box 6188, 99000 Beit Shemesh (IL).

(74) Agent: **REINHOLD COHN AND PARTNERS**; P.O. Box 4060, 61040 Tel-Aviv (IL).

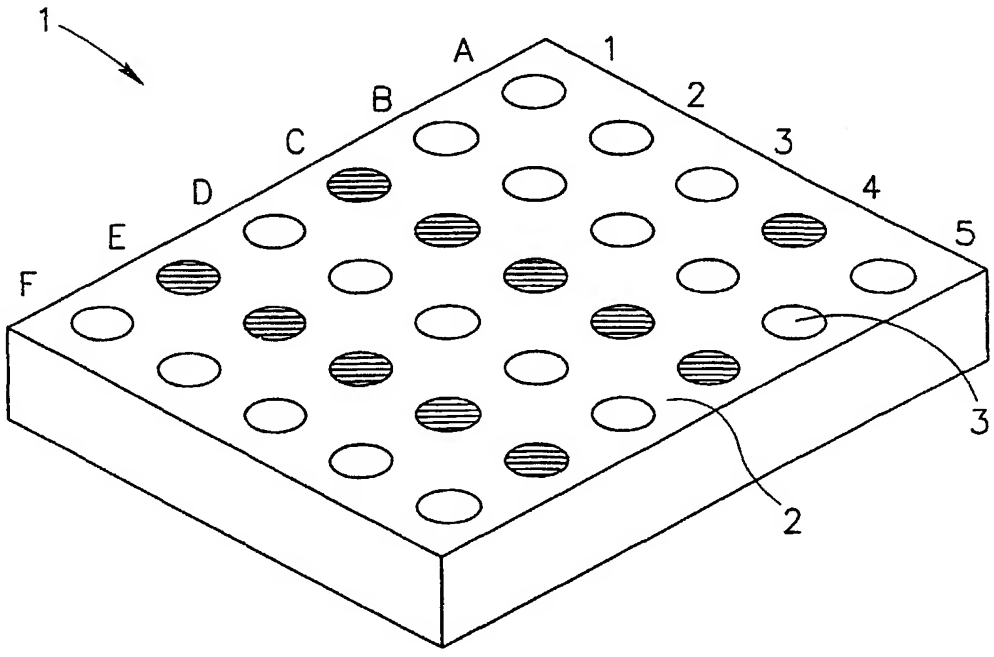
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ARRAYS OF NUCLEIC ACID SEQUENCES



(57) Abstract: The invention concerns an array of catalytic nucleic acid sequences or sequences having a potential of becoming catalytic having a predetermined spatial arrangement. The array may be used for detection purposes or for the purpose of *in vitro* evolution.

WO 01/00869 A2

ARRAYS OF NUCLEIC ACID SEQUENCES

FIELD OF THE INVENTION

The present invention concerns arrays of nucleic acid sequences and methods of detection using said arrays. The present invention further concerns a method for detecting an analyte in cells, as well as a method for gene therapy.

BACKGROUND OF THE INVENTION

Arrays of different nucleic acid sequences, wherein the sequences have a predefined spatial arrangement so that the location of each sequence is known, are used for hybridization-based detection purposes. These arrays, contain a plurality of different oligonucleotides, which are complementary, to some degree, to oligonucleotides which are to be detected in a sample. Then, the array is brought into contact with the sample under conditions allowing hybridization of nucleic acid sequences, and the presence of hybridization complexes is detected, for example, by using fluorescent labels. Typically, the arrays of oligonucleotides are present as *"biochip"*, i.e. an array of oligonucleotides arranged in a very small area. Using developed robotic techniques for micro-patterning of biological molecules, the predetermined arrangement of oligonucleotide sequences can be packed to a single biochip the size of an electric microchip which may be read and analyzed by a computer.

The nucleic acid hybridization-based arrays, are used for a plurality of detection purposes, such as multiplexing detection of many different sequences in a sample, sequencing of longer sequences by hybridization (SBH), etc.

However, all nucleic acid sequences arrays, are based on probes for hybridization to complementary sequences.

Catalytic nucleic acid sequences, notably ribozymes, are molecules, typically RNA molecules, which have enzyme-like catalytic activities associated with cleavage, splicing or ligation of nucleic acid sequences. The typical substrates for ribozymes catalytic activities are RNA molecules, although ribozymes may catalyze reactions in which DNA molecules (or even amino acids) serve as substrates.

It has recently been proposed to use ribozymes in order to treat diseases or genetic disorders by cleaving a target RNA, such as viral RNA or messenger RNA transcribed from genes that should be turned off. This was proposed as an alternative to blockage of the RNA transcript by the use of antisense sequences. Owing to the catalytic nature of the ribozyme, a single ribozyme molecule cleaves many molecules of target RNA and therefore therapeutic activity is achieved in relatively lower concentrations than those required for antisense treatment (WO 96/23569). However, in all cases, the ribozyme acts directly on the viral RNA directly.

The use of ribozymes for diagnostic purposes has been only seldomly mentioned. WO 94/13833 describes a method for detecting nucleic acid molecules in a solution by tailoring a specific ribozyme molecule having two regions, one complementary to the nucleic acid sequence to be detected, and the other complementary to a co-target molecule bearing a detectable label. The ribozyme is able to specifically and reversibly bind both to a selected target nucleic acid sequence and to the labelled co-target. When both the target and the co-target are bound, the ribozyme undergoes a conformational change which renders it active, and in its active condition it is able to cleave the label of the co-target, and the free label can then be detected. Upon cleavage of the co-target, the ribozyme is able to re-associate with an additional co-target, cleaving more label and producing more detectable signals.

WO 98/08974 concerns a nucleic acid molecule termed *"proto-nucleozyme"* that has *a priori* no catalytic activity but which becomes catalytically

active by formation of a complex with a co-factor to form a catalytically active complex termed *"nucleozyme"*.

A proto-nucleozyme is in fact a nucleozyme with a missing component, which missing component is completed by the co-factor. The complex between the proto-nucleozyme and the co-factor may also at times be referred to as *"catalytic complex"* and is in fact a nucleozyme since it has catalytic activity. The proto-nucleozyme may consist of deoxynucleotides (dNTP's), ribo-nucleotides (rNTP's), as well as other nucleotides such as 2′–O-methyl nucleotides, or any combinations of these.

GLOSSARY

The following terms may be used at times throughout the specification:

Oligonucleotide – a sequence of nucleotides. May be composed entirely of dNTP's, rNTP's or a combination of both, and may comprise non-naturally occurring nucleotides such as 2′-O′-methyl nucleotides or a combination of the above.

Oligonucleotide species – a plurality of oligonucleotide molecules featuring essentially the same nucleotide sequence. Each oligonucleotide specie has a different sequence different from that of another oligonucleotide sequence.

Catalytic oligonucleotide – a nucleic acid sequence featuring an enzyme-like activity, such as cleavage, ligation, splicing-out (cleaving both ends of a short nucleic acid sequence to remove it from a longer sequence and ligating the ends of the cut), splcing-in (cleaving open a nucleic acid sequence, inserting another short nucleic acid sequence and ligating the ends of the cut), rearrangement, as well as additional catalytic activities such as phosphorylation, kinase like activity, addition or removal of other chemical moieties, biotinilation, gap filling of missing nucleotides, polymerization, etc.

Pre-catalytic oligonucleotides – oligonucleotides which are initially inactive but can become catalytically active in the presence of a suitable molecule which serves as their co-factor, an example of such oligonucleotides are allosteric ribozymes as featured in U.S. 5,589,332 and WO 98/0897, WO 94/13833. The term is equivalent also to the term *"allosteric oligonucleotide"* as will be explained below.

Co-factor – a molecule or a moiety within a molecule that reverts the pre-catalytic oligonucleotides (*"allosteric oligonucleotide"*) from their initially inactive form to an active form. This may be a protein, a peptide, a hormone (or a moiety therein), a nucleic acid sequence (mRNA, DNA), a nucleotide, a toxin or any other molecule or moiety from a biological source. Typically it is a cellular or viral protein or mRNA which characterizes a *"desired cell population"* (see below).

Candidates for catalytic oligonucleotides – oligonucleotides which are candidates for evolving, by *in vitro* evolution, to catalytic oligonculetodies.

Array – a predetermined spatial arrangement of catalytic oligonucleotides, pre-catalytic oligonuceotides or candidates for catalytic oligonucleotides present on a *solid support,* (see below) or in *an array of containers* (see below) wherein each *oligonucleotide-bearing location* (see below) or each container, respectively, holds one oligonucleotide species.

Oligonucleotide-bearing location – a specific area in the solid support on which there are held oligonucleotides of the same specie. The solid support may hold several identical such locations, for example, 100 dots all having oligonucleotide specie "X", 100 dots having nucleotide species "Y", etc.

Solid support – a rigid surface which carries the array of catalytic oligonucleotides, pre-catalytic nucleotides or candidates for catalytic oligonucleotides.

Array of containers – an arrangement of liquid-holding containers, for example, a multi-well arrangement, wherein each container holds a single species of oligonucleotides.

Positive selection - a step of *in vitro* evolutions wherein oligonucleotides which show cataytic activity under a desired condition are selected.

Negative selection – a step of *in vitro* evolution wherein oligonucleotides which show catalytic activity under a non-desired condition or which do not show catalytic activity under a desired condition are discarded.

Heterologous nucleic acid sequence (also referred to at times as "heterologous gene/heterologous coding nucleic acid sequence) – a sequence which has a potential of being expressed to give a *"desired product"* (see below). The sequence is introduced from the outside into the cell by use of a vector. The sequence may be a sequence from another species, a sequence which is the correct version of a mutated and faulty sequence already present in the cell (which mutated sequence causes production of a mutated protein), or sequence already present in the cell which is introduced to the cell in order to raise its level in the cell. In accordance with the invention the expression of said heterologous sequence is regulated (turn on or off) by an *"allosteric oligonucleotide"* (see below).

Allosteric oligonucleotide – this term refers in fact to pre-catalytic oligonucleotides, which are initially inactive but become catalytically active in

the presence of a suitable molecule which serves as the co-factor. Typically the co-factor of the allosteric oligonucleotide is a cellular or viral protein or mRNA.

Desired product – the protein or mRNA produced by the expression of the heterologous nucleic acid sequence. Usually is either a detectable label or has a therapeutic effect.

Desired cell population – a specific group of cells in which a detection or therapeutic action is to take place. The population is characterized in that cells of the population contain a co-factor for an allosteric oligonucleotide, while other cell populations (not *"desired"*) do not contain the co-factor. The desired cell population may be a viral infected cell population, cancer cells, cells of a specific tissue or cells of a specific stem cell origin, etc.

SUMMARY OF THE INVENTION

The present invention concerns an array of catalytic oligonucleotides, pre-catalytic oligonucleotides or candidates for catalytic oligonucleotides immobilized on a solid support, wherein each oligonucleotide-bearing location of the solid support, bears essentially a single species of said oligonucleotides.

The array of oligonucleotides may include oligonucleotides which are *a priori* active, such as state of the art ribozymes (made from rNTP's, nNTP's or a combination of rNTP's and dNTP's) which can feature their catalytic activity once their proper substrate is supplied. Alternatively, the array may include pre-catalytic nucleic acid sequences, i.e. oligonucleotides which are *a priori* inactive, but have the potential of becoming catalytically active, for example, in the presence, of a co-factor (*"effector"*) which can be a molecule which presence is to be assayed in a sample.

By a third alternative, the array may contain a huge number of synthesized oligonucleotides, having a random or semi-random sequences, which oligonucleotides are candidates for evolving into catalytically active oligonucleotides in the various *in vitro* evolutions techniques.

In all three cases (i.e. catalytic oligonucleotides, pre-catalytic oligonucleotides, and oligonucleotide candidates for catalytic activity) use is made of the predefined spatial of arrangement of the oligonucleotides in the array wherein all oligonucleotides present in a single location or in a single well are of the same species. Thus, once it is possible to determine, as will be explained hereinbelow, in which location of the array catalytic activity took place, it is possible immediately to know what is the sequence of the oligonucleotide present in said location or how to obtain oligonucleotides of same sequences (for example, by knowing the precise clone which produced them). This information may be used in one of two ways: according to the *"detection aspect of the invention"* catalytic activity of a specific catalytic, or pre-catalytic oligonucleotide species is expected only in the presence of a specific assayed molecule. The assayed molecule may be the substrate of a *trans* acting, specific catalytic nucleic acid sequence, or a co-factor of a specific pre-catalytic molecule. Thus, by knowing which specific location featured catalytic activity, it is possible to determine which oligonucleotide species was present in that location, and thus to know which assayed molecule (i.e. said substrate of catalytic oligonucleotide or co-factor of said specific pre-catalytic oligonucleotide) was present in the sample.

In another variation of said invention, a proto-nucleozyme with a known co-factor requirement is mutagenized and placed on a chip so as to be increase specificity and give more information about unknown other analyte. The randomized pool of proto-nucleozymes that are received after relatively few rounds of evolution are cloned. The pattern of different proto-nucleozymes to which is seen on the chip would indicate more precisely the analyte which activate and might make it possible to characterize the unknown analytes into various *"families"* and give general prediction of their structure.

In accordance with the *"in vitro evolution aspect of the invention"* the precise location of each synthesized oligonucleotide, which is a candidate for evolving to a nucleic acid sequence having catalytic activity, in the array is known or alternatively it is known which clone produced the candidate sequence. Then by

determining which locations bear catalytically active oligonucleotides under a desired condition (positive selection step), it is possible to conclude what was the sequence of oligonucleotides in that location and thus to select only oligonucleotides which are catalytically active under said condition. Alternatively, subsequently or simultaneously, it is possible to determine which locations featured catalytic activity in the presence of a non- desired condition, or which did not exhibit catalytic activity under the desired condition and these locations are not selected (negative selection step).

The present invention also concerns an array of catalytic oligonucleotides, pre-catalytic oligonucleotides, or candidates for catalytic oligonucleotides present in an array of liquid-holding containers, wherein each container holds essentially a single species of said oligonucleotides.

The array of containers, may be for example an array of wells, for example a 96-well arrangement. In each well, there is present essentially one single species of catalytic oligonucleotide, pre-catalytic oligonucleotides or candidates for catalytic oligonucleotides as explained above, and it is *a priori* known, what is the sequence present in each well. Then, the detection as explained in *"the detection aspect"* of the invention or the *in vitro* evolution as explained in the *"in vitro evolution aspect"* above, can take place, wherein one container is equivalent to one *"oligonucleotide bearing location"* in the solid support.

The term *"immobilized"* refers hereinafter to any type of binding, both covalent and non-covalent of the oligonucleotides to the solid support as well as to absorption or adsorption of the nucleic acid sequences to the solid support. The binding may be covalent by using active groups such as NH_2. The binding may also be carried out by using nucleic acid binding proteins, for example DNA binding proteins which have been modified to react with the solid substrate. The nucleic acid binding proteins may be derivated for example by thiols so they can bind to glass. Other binding methods of oligonucleotides are well known in the art.

The solid support may be of two types: a *"static solid support"* and the other *"retrievable solid support"*.

The *"static solid support"* refers to a solid support wherein once the nucleic acid molecules are immobilized thereon, they essentially remain in that same position indefinitely. This means that each *oligonucleotide-bearing location* in solid support stays in its own place. The static solid support is mostly used in the *"detection aspect of the invention"* (see below), but can also be used in the *"in vitro evolution aspect of the invention"*.

The term *"retrievable solid support"* refers to types of solid support wherein *nucleotide-bearing locations* of the solid support can independently be separated from other nucleotide-bearing locations. Examples of such solid supports are beads, such as sepharose beads or magnetic particles wherein a specific location of beads can be picked up and transferred to a new surface or to a new vessel for further processing. The *"retrievable solid support"* is mostly used for the *"in vitro evolution aspect of the invention"*.

In each oligonucleotide-bearing location there is immobilized essentially a single species of oligonucleotides. However, it is possible that a number of identical oligonucleotide species are immobilized in several different adjacent locations in the array. The fact that there exists a plurality of locations all bearing the same oligonucleotide species enables to carry out various statistic manipulations as will be explained hereinbelow.

The array of the present invention may be formed as a *"biochip"* i.e. an array of biological molecules present in a very small area.

Using developed robotic techniques for micro-patterning of biological molecules, it is possible to create a predetermined arrangement of biological molecules containing thousands of different molecules packed onto a single biochip the size of an electronic microchip which can be read automatically by a computer. Typically, the biochip consists of a small glass plate coated with a gel. Using a special robotic device the array of oligonucleotides may be placed on the gel. The biochip can be analyzed for a detectable signal, for example fluorescence, and the result interpreted as will be explained hereinbelow, wherein each location of oligonucleotides present on the biochip is precisely known. Typically, a biochip is

used for the simultaneous detection of a plurality of different analytes, or for *in vitro* evolution purposes.

The present invention further concerns an assembly for the detection of a location of catalytic activity in the array of the invention. Typically, the assembly comprises the array of the invention, as well as signal producers, which are capable of producing a detectable signal upon the activation (catalytic activity) of the catalytic oligonucleotides. Such signal producers may be an integral part of the oligonucleotides of the array themselves. For example, where the catalytic oligonucleotide undergoes an auto-catalytic activity (for example, cleavage or splicing) the oligonucleotides of the array may comprise a fluorescent moiety which becomes detached upon cleavage or splicing and the location where catalytic activity took place can be detected by lack or decrease of fluorescence emission therefrom.

Alternatively, the signal producers means may be molecules independent from the oligonucleotides of the array. For example, where the catalytic activity of the oligonucleotides is ligation, those oligonucleotides which are catalytically active may ligate to themselves a short sequence being a fluorescent moiety, the short sequence in that case is the *"signal producing means"*. Thus, only fluorescent bearing locations are considered as holding catalytically active oligonucleotides. In such a case, the assembly comprises the array as well as said label bearing sequence which can be attached by ligation.

Labeling can be also by a two-step procedure in which catalytic activity (cleavage, ligation, splicing, phosphorylation) creates a site to which a detectable label can be attached in a subsequent step, (for example, by use of antibodies, by use of complementary sequences bearing labels, etc.). In that case the assembly contains said detectable label as well as elements for its attachment to the oligonucleotides.

The detection may be read optically by a fluorescent label, or by a dye producing a color reaction as known in the art or by other techniques known for detection.

Another possibility is the use of magnetic particles as labels, which can be detected by devices capable of detecting weak magnetic fields.

The present invention further concerns a system for the detection of the location of catalytically active oligonucleotide species in an array, comprising the above assembly as well as a detector capable of reading said signal. Such a detector where the signal is optical, may be for example a CCD camera attached to an analog-to-digital converter, and to a computer capable of analyzing the signal and indicating specifically in which location a catalytic reaction took place. Where the label is magnetic, the device should be a magnetic-field sensor capable of detecting precisely the presence and location of weak magnetic fields.

By one aspect termed *"the detection aspect of the invention"* the oligonucleotide species, which may be catalytic or pre-catalytic, are used to detect the presence of an assayed molecule (also referred at times as *"analyte"*) in a sample. The analyte may be a separate molecule (protein, nucleotide molecule) or a moiety within a large molecule (an epitope of a protein, a short sequence of nucleotides within a longer sequence). The oligonucleotides of the array may be *a priori* catalytically active, for example, having *trans* enzymatic activity, and the assayed molecule (analyte) can then be the substrate on which they are active. In that case, if catalytic activity is cleavage or splicing, the assayed molecule can be labeled prior to carrying the assay, for example, by attaching to it a fluorescent label. In this example the *trans* activity produces a detectable signal. Where the catalytic activity is ligation of fluorescent-bearing sequence, there may be no need to label the assayed molecule (analyte) and all that is needed is to provide a label-carrying sequence which may be attached to the substrate by ligation. The catalytically active molecule may be essentially naturally occurring ribozymes that act in *trans*, or alter molecules that are tailored, or *in vitro* evolution evolved sequences, developed so that their substrate is the desired analyte.

However, by a preferred embodiment the oligonucleotides are *a priori* non-active (i.e. they are pre-catalytic oligonucleotides), and become catalytically active only in the presence of the assayed molecule, which serves as their co-factor.

WO 98/08974, incorporated herein by reference discloses such pre-catalytic oligonucleotides which are termed therein *"proto-nucleozymes"*. Preferably, such proto-nucleotzymes should be of the type that they lack a missing component essential for catalytic activity (such as a segment in their core region, or in their stem region) and hence are inactive. A molecule termed *"co-factor"* can complete the missing region and thus cause the proto-nucleozyme to become active. Only if the co-factor (being the assayed molecule, i.e. the analyte) is present, these proto-nucleozymes revert to their active form and produce a detectable signal. The co-factor, i.e., may be any type of molecule, such as a nucleic acid sequence, a protein, a hormone, a drug, a nucleotide, etc., which presence is to be detected in a sample as explained in detail in WO 98/0897.

In accordance with the detection embodiment of the invention, an array of catalytic oligonucleotide or pre-catalytic oligonucleotides can be prepared, wherein each location in the array bears one species of oligonucleoties. The different species of the oligonucleotides may become catalytically active in the presence of different co-factors, for example, different proteins. Alternatively, each species of oligonucleotides may become catalytically active in the presence of different epitopes of the same protein. Thus by determining which species of oligonucleotides become catalytically active, it is possible to do *"epitope mapping"*. This manner enables to detect, with a very high specificity, a specific protein since all of its specific epitopes are characterized. Only when a correct combination of different oligonucleotide species (present in a correct combination of locations in the array) show catalytic activity, it indicates that a protein having all the specific epitopes was present in the sample. Such a detection may eliminate false negative results, usually encountered in antibody-based detection assays due to cross-reactivity of antibodies with several proteins which have some common epitopes.

It is of course possible to combine the two possibilities, i.e. to create an array which detects in a specific manner plurality of different proteins, by detecting a combination of several epitopes for each of these proteins.

Thus, the present invention concerns a method for detecting the presence of at least one co-factor in a sample comprising:

(i) providing an array of the invention comprising pre-catalytic oligonucleotides;

(ii) incubating the array with a sample under conditions allowing catalytic activity of oligonucleotides;

(iii) determining, by using the assembly of the invention, in which locations catalytic activity took place, said determination indicating the specific oligonucleotide species which became catalytically active, thereby indicating the presence in the sample of the specific co-factor which activates the specific species.

The co-factor as disclosed above is of course the assayed molecule, or analyte as defined previously.

By another aspect of the invention termed *"the in vitro evolution aspect of the invention"* the array of the invention is used for *in vitro* evolution of catalytic oligonucleotides.

Typically, during *in vitro* evolution, catalytically active oligonucleotides, (for example, ribozymes) are separated from catalytically inactive oligonucleotides, by separating between immobilized and diffusable oligonucleotides. The supernatants may contain oligonucleotides which have undegone catalytic activity, for example, oligonucleotides which are freed from an immobilized solid support by auto-catalytic cleavage. Thus by collecting only the supernatant only catalytically active oligonucletides are collected. Alternatively, the supernatant may hold oligonucleotides which are not catalytically active, for example, oligonucleotides which did not become immobilized, through ligation, to the solid support. The catalytically active oligonucleotides may be obtained in a positive selection step, (i.e. obtaining oligonucleotides having a catalytic activity under a desired selective condition), or discarded in a negative selection step (i.e. discarding oligonucleotides which have catalytic activity under a non-desired selective condition).

Thus, usually *in vitro* evolution methods contain lengthy procedures of separating supernatants from immobilized ribozymes, for examples, by various centrifugation techniques followed by extensive rinsing.

In accordance with the *in vitro* evolution embodiment of the present invention, it is possible to produce an array of candidates for catalytic oligonucleotides, then expose said candidates to a specific selective condition (which may be a desired or non-desired selective condition), and then select, (by one of the two manners disclosed hereinbelow) only oligonucleotide-bearing locations which hold oligonucleotides that have undergone catalytic activity under said condition. Said oligonucleotides may be obtained for the positive selection step, (if the conditions to which the array is subjected is desired) or discarded in a negative selection step (if the condition to which the array is subjected is not desired).

By one embodiment of the *in vitro* aspect of the invention obtaining or discarding the selected oligonucletoides is done physically, i.e. by collecting from a retrievable support only the locations of catalytically active oligonucleotides.

According to this embodiment of the invention, an array of a plurality of different oligonucleotide species, which are candidates for evolving to catalytic oligonucleotides having a desired catalytic activity under a specific condition, are prepared, and subjected to a specific condition under which they should feature the catalytic activity. The condition may be the presence of a co-factor, a specific temperature, a specific pH, ionic strength, etc. The oligonucleotids are allowed to undergo catalytic activity. Those locations of the array bearing oligonucleotides which underwent catalytic activity are identified by using the assembly of the invention, leading to production of a detactable signal, and the system of the invention reading said signal as described above.

Then, those locations bearing said catalytically active species are retrieved, for example, by collecting the sepharose beads, or magnetic particles on which the oligonucleotides are immobilized. The sequences are separated from the obtained locations and are transferred to a nucleic acid analyzer for determining the

sequence of the oligonucleotides. Alternatively, the oligonucleotides may be processed for reconstruction in order to revert the oligonucleotides which underwent auto-catalytic activity back to the sequence of original catalytic oligonucleotides prior to auto-catalytic activity. Where the catalytic activity is cleavage said reconstruction may be for example by reconstructing the cleaved sequence. After reconstruction the oligonucleotides are transferred to a new solid substrate, in order to create a new array containing only the selected oligonucleotides. Then, the new array can be subjected again to a specific condition, which may be again a positive selection step albeit under stricter criteria (higher temperature, lower concentration of co-factor, etc.) or a negative selection step. Again, only desired locations are collected, processed for reconstruction, so that the sequence that was changed due to auto-catalytic activity is produced again, and transferred to create a new array.

Using the method of the invention, separation between catalytically active and catalytically inactive oligonucleotides may be carried out without separation of the supernatants and the immobilized portions in the reaction vessel, but rather simply by physically picking up those locations which are desired, i.e. locations which are labeled.

By using this first embodiment, there is no need to know *a priori* what was the exact sequence of the oligonucleotide immobilized in each location on the array, as those sequences of oligonucleotides which are catalytically active are physically collected. Thus, it is possible to create, by a oligonucleotide synthesizer, a plurality of different sequences and deposit these sequences in different locations, and all that it is important is to ensure that on each location only one species of sequences is deposited. Alternatively, it is possible to create by genetic engineering techniques a library of different nucleic acid sequences and ensure that in each location the sequences of the oligonucleotides are obtained from a single clone. After several rounds of selection, wherein in each round only some locations are collected, there remain a small number of locations which bear oligonucleotides having a desired activity (for example, catalytic activity in the presence of a

specific co-factor; catalytic activity in the presence of a specific temperature, etc.). Then, and only then, the sequences of those active oligonucleotides may be determined, for example, by automatic nucleic acid sequencer.

By a second embodiment of the *in vitro* evolution aspect of the invention, it is *a priori* known what is the sequence of candidate oligonucleotides present in each location. This is possible either by depositing on each specific location a known oligonucleotide species (for example produced by a nucleic acid synthesizer), by synthesizing *in situ* on the solid support different species in different locations, or by placing in each location a sequence obtained from one tranfected clone prepared by genetic engineering, wherein it is *a priori* known which sequence is present in which clone. Then, after exposing the array to a set of selected conditions, it is not necessary to physically obtain those locations which carry catalytically active oligonucleotides, and all that is necessary is to determine, based on the computer's memory, what is the sequence deposited, or synthesized, in the location which showed catalytic activity, or to determine from which clone it was obtained.

Then these selected sequences may be synthesized again for mass production, or the clones producing these sequences can be identified and used for mass production. Alternatively, if it is desired to continue the selection process only the selected sequences may be synthesized again (either separately and then deposited or *in situ* on the solid support) in order to create a new array having only oligonucleotides which have shown a specific catalytic activity under a desired condition. Then the new array may again be subjected either to a second selective condition, or to the first condition under more strict criteria, and again, by using the computer's memory, it is possible to know what is the sequence of the oligonucleotide positioned on those locations which showed a detectable signal indicating catalytic activity. If genetic engineering methods are used to create the plurality of sequences, it is not necessary to store in the computer the explicit sequence of oligonucleotides present in each location, but rather only store the identification of the clone from which the sequence in that location was obtained.

Then when a desired location featuring the desired activity is selected, it is possible to revert to the original clone and determine the sequence of said clone.

Use of the array of the invention, wherein each location bears a sequence having a computer stored sequence, enables the use as candidate nucleotides for evolving in the *in vitro* evolution methods, sequences bearing non-naturally occurring nucleotides.

In state-of-the-art *in vitro* evolution methods, selected oligonucleotides are amplified utilizing enzyme-based amplification methods such as PCR. By such methods it is possible to amplify only sequences composed solely of naturally-occurring nucleotides. However, in accordance with the present invention, it is possible to produce, by a nucleic acid synthesizer, sequences which contain, or are composed of, non-naturally occurring nucleotides (such as O-methyl nucleotides). Then each species of sequences is placed in a single location and the sequence of each location is known (stored in the computer's memory). When it is desired to amplify sequences which showed catalytic activity in a specific location, all one has to do is determine, utilizing the computer's memory, which sequence is present in said location and then synthesize again this sequence on the synthesizer. The sequence itself is not used as a *"template"* in enzymatic-based amplification methods. Thus it is possible to select and amplify desired sequences containing non-naturally occurring nucleotides.

By another embodiment of the *in vitro* aspect of the invention, a plurality of identical arrays of the invention is used, wherein in each location of the identical arrays, there is present essentially a single species of oligonucleotides which are candidates for catalytic activity. The plurality of the arrays are exposed to a desired selection condition.

Then a *"cut-off percentage"* is set. For example, if a specific location showed catalytic activity under the selected condition in at least 90% of the arrays, then the specific location is defined as one carrying catalytically active oligonucleotides which should be obtained. Alternatively, or in addition, a corresponding plurality of arrays is exposed to a non-desired condition, i.e. a

condition wherein no catalytic activity should take place. Again a *"cut-off percentage"* is selected, for example, 1%. Only these locations which showed catalytic activity of less than 1% of the arrays (i.e. less than 1%), under the non-desired condition, and 90% catalytic activity under the desired condition are chosen. Alternatively, it is possible to simply choose the best location (which had the highest activity under a desired condition and the lowest activity under a non-desired condition), without, *a priori*, defining *"cut-off"* levels.

Usually, under state of the art *in vitro* evolution methods, the selective conditions are initially quite permissive but they become more stringent as selection cycles proceed. However, in accordance with the present invention, it is preferable to start the selection process by first exposing the oligonucleotides to the most stringent conditions desired. For example, where it is desired to select for oligonucleotides which are catalytically active only in the presence of a specific co-factor, it is preferable to first expose the plurality of the arrays to the lowest concentration of co-factors desired. If no catalytic activity is detected in none of the locations, then the concentration can gradually rise, until a concentration is reached wherein a minimal detectable signal is produced. Another example is to choose catalytic oligonucleotides which are active in the lowest temperature possible. Again, it is possible to expose nucleotides to a very low temperature and if no catalytic activity is evident, slowly raise the catalytic activity.

By yet another aspect, the present invention concerns a method for *in vivo* or *in situ* determination of an analyte (or assayed molecule) in cells. The analyte can be any compound whose presence is to be detected inside cells, such as a protein, a metabolite, a hormone, a drug, a nucleoside, molecule from an infectious agent, etc. In accordance with the method, pre-catalytic oligonucleotides of the type that the analyte serves as a co-factor for its catalytic activity, according to the teaching of WO 98/08974 are prepared, in such a manner that its auto-catalytic activity changes a detectable property in the oligonucleotide. For example, the pre-catalytic oligonucleotide may comprise a fluorescent label, as well as a moiety capable of quenching fluorescence. If the fluorescent label and the quencher are adjacent to

each other, the fluorescence is quenched, and essentially there is no fluorescent emission. However, in the presence of an appropriate co-factor, the pre-catalytic oligonucleotide becomes catalytic, and cleaves the fluorescent label, which by diffusion is distanced from the quencher so its fluorescence radiation is evident. Detection of fluorescence indicates the presence of a catalytic activity of the oligonucleotide, which in turn indicates the presence of the analyte which served as the co-factor of the pre-catalytic oligonucleotide. This construct is present inside a vector capable of entering cells, such as a plasmid, a suitable viral vector, etc.

Once the construct enters the cell, the pre-catalytic nucleic acid sequence can become catalytically active if the analyte is present inside the cell. In such a case, the label will be distanced from the quencher, and the cell will show fluorescence, which fluorescence may be determined by any manner known in the art.

Thus the present invention concerns an insertion vector for inserting at least one nucleic acid construct into a cell; the construct comprising a protonucleozyme being a nucleic acid molecule or complex of nucleic acid molecules having essentially no catalytic activity but can complex with a co-factor to form a nucleozyme which possess catalytic activity, the protonucleozyme lacks a component essential for the catalytic activity of the nucleozyme and said co-factor provides said component; the nucleic acid construct comprising to a detectable moiety which changes at least one detectable property upon *cis* catalytic activity of the nucleozyme, the co-factor of the protonucleozyme being a molecule or a moiety within a molecule found in cells.

The present invention further concerns a method for determining the presence of an analyte *in vivo* or *in situ*, a cell comprising:

(i) utilizing the vector of the invention to insert the nucleic acid molecule construct to the cell, wherein the analyte to be detected is the co-factor of the protonucleozyme; and

(ii) determining the existence of the detectable labels in the cells, said existence signifying the presence of the analyte in a cell.

The method and construct of the invention are suitable for determining, or monitoring the presence of a single analyte (protein, metabolite, hormone drug, toxic substance, etc.) in the cells *in vivo*, or *in situ*, and may be used in cases where it is desired to *"view"* the inside of the cell, for example, to monitor the spread of a viral infection, the penetration of a pharmaceutical agent, etc.

Alternatively, the method, and the construct may be used to determine, or monitor a plurality of different analytes in the cell simultaneously.

In such a case, the vector will contain a plurality of pre-catalytic sequences, each one capable of being activated by a different analyte which serves as its specific co-factor.

In the presence of the suitable analyte, the auto-catalytic activity is carried out, and a part of the substrate, bearing the fluorescent label, is distanced from the quencher and thus fluorescence is emitted. In accordance with the plurality embodiment, each cleaved part of the substrate carries also a *tag* sequence enabling its identification.

For example, the *tag* may be an immunogenic moiety, capable of being recognized by specific antibodies. According to a preferred embodiment, however, the *tag* sequence, which identifies the type of the pre-catalytic nucleic acid attached thereto (and by this identified which analyte served as a specific co-factor for this pre-catalytic sequence) is a nucleic acid sequence. The *tag* sequence can be detected by breaking up cells, so that their contents are released to the medium. Then, the contents of the released cells are brought into contact with an oligonucleotide chip, containing sequences complementary to said *tag* sequences. By knowing precisely in which location on the chip, the complementary sequence is placed, and by knowing on which location fluorescence is present (due to hybridization of the fluorescent-label carrying tag sequence) it is possible to know which analytes were present in the sample.

Thus the present invention concerns a vector as described above comprising a plurality of nucleic acid constructs, each construct comprising a protonucleozyme which becomes catalytically active through complexing with a different co-factor,

each nucleic acid construct further comprising an identification *tag* sequence, said identical tag sequence being adjacent to the detectable label, and being detached, together with said label, from the construct through catalytic activity.

The invention further concerns a method for the simultaneous determination of a plurality of analytes *in vivo* or *in situ* in a cell, the method comprising:

(i) utilizing the vector of the invention comprising a plurality of different constructs to insert the constructs into the cell wherein each analyte to be detected is the co-factor of a different protonucleozyme;

(ii) processing the cell to liberate its contents into a medium;

(iii) identifying in the medium the nature of detached, label-bearing identity tags thereby determining the identity of the plurality of analytes in the cells.

By yet another aspect, the present invention concerns a method for the control of expression of a heterologously induced nucleic acid coding sequence. The term *"heterologously induced nucleic acid sequence"* refers to a nucleic acid sequence, which is induced to the cell from an external source, and has a potential of becoming expressed to give a desired product. The term *"heterologous"* should be understood to refer both to a sequence coding for a protein which the cell originally does not have at all (for example from another species), a sequence which has a mutated counterpart in the cell, and a correct sequence is heterologously introduced (for example to compensate for lack of metabolite) as well as a sequence which is already present in the cell, and it is desired to increase the level of mRNA of said sequence.

According to the method of the invention, said heterologous sequence is placed in a nucleic acid construct, in such a manner that, only upon activity of a catalytic nucleic acid sequence it can either revert from non-expression to expression (up-regulated) or from expression to non-expression (down-regulated).

The term *"expression"* refers to the fact that the sequence produces the desired product (protein or mRNA). Production of non-desired products (for

example, fragments of the full protein) or non-production of all are collectively referred to as *"non expression"*.

For up-regulation purposes the heterologous nucleic acid sequence may be broken into two parts with an intervening sequence thereinbetween. Only splicing out of this intervening sequence will enable expression of the sequence to produce the desired product (without said splicing only a short fragment is produced). The intervening sequence may be a substrate for splicing activity of a catalytic nucleic acid sequence *in trans*, or alternatively, may be the catalytic nucleic acid sequence itself, which by *cis* - auto catalytic splicing, excising itself out of the heterologous sequence and ligating the two open ends thus enabling production of the desired product.

By another alternative of up-regulation, the heterologous nucleic acid sequence may, *a priori*, comprise an inhibitory control sequence. Only cleaving or splicing out of this inhibitory sequence causes the catalytic nucleic acid sequence to be expressed. An example of down-regulation is a heterologous nucleic acid sequence that, *a priori*, is expressed to give the desired product. Only upon catalytic activity, such as cleaving or splicing, its expression is terminated, for example, due to cleaving of a promoter, or due to the fact that cleaving cutting the sequence to give non-desired short fragments.

The catalytic oligonucleotide sequence should be of the allosteric type, i.e. of a type which is initially inactive, but which becomes catalytically active in the presence of a co-factor, for example, as specified in U.S. 5,589,332 or WO 94/13833 activated by a co-factor. A preferred embodiment is the proto-nucleozymes specified in WO 98/08975 which lacks a component essential for its catalytic activity and said co-factor completes said component.

The co-factor is a molecule or moiety within a molecule which is present essentially only in the desired cell population and is not present in other cell populations.

Thus, in accordance with the method of this aspect, only in the presence of a co-factor, does the allosteric oligonucleotide become active, in turn regulating (up- or down-regulation) the expression of the heterologous sequence.

By one embodiment, the heterologous nucleic acid sequence is intended to produce a detectable label for diagnostic purposes.

For example, the heterologous sequence may be the luciferase gene which, when expressed, produces a luciferase enzyme capable of providing bioluminescence. If the luciferase gene is introduced into the cell with an allosteric oligonucleotide present in such a manner which it cannot be expressed (for example the allosteric oligonucleotide interrupts coding sequence) then only if the correct co-factor of the allosteric oligonucleotide is present, then the catalytic sequence becomes active, splices itself out of the luciferase gene sequence, thus enabling production of bioluminescence. If the co-factor is for example an analyte to be detected inside cells, such as a viral protein, bioluminescence can be indicative of the presence of the viral protein within the cells. This fact may be used in screening for drugs wherein it is desired to detect those cells which respond to the drug by change of a cellular component which is the co-factor.

In an analogous manner, the heterologous sequence may be a sequence of a gene used in gene therapy for therapeutic purposes. In such a case, the allosteric oligonucleotide can be used to turn on or off production of a therapeutic product (being the desired product) in response to the presence of a co-factor present inside a cell. In gene therapy, it is often advantageous that the expression of the heterologous gene is controlled so as to be expressed only in a specific sub-set of cells (being the desired cell population). For example, where the heterologous gene is a cytotoxic gene such as a gene coding for a toxin or caspase, it is advantageous that it would be expressed only in viral infected cells (such as in the case of AIDS) or cancer cells so as to selectively destroy only these cells. If the cytotoxic gene expression is regulated by an allosteric oligonucleotide, then cell death by expression of the cytotoxic sequence can be regulated, so that only viral infected or cancerous cells, which produce the co-factor required for the activity of the

allosteric oligonucleotides are destroyed, while cells not comprising said co-factor are not destroyed.

The heterologous sequence used for therapeutic purposes can be sequences coding for cytotoxic agents or any other sequence which expression has a beneficial effect in cells.

For example, in many cases genetic diseases cause aberrant mRNA production (leading to aberrant protein production), which may be expressed only in certain cell types such as liver. If the aberrant mRNA or the aberrant protein is used as the co-factor of the allosteric oligonucleotide, and the heterologous gene which is regulated by the allosteric oligonucleotide is the correct sequence, then the aberrant mRNA (or the aberrant protein) causes production of the correct mRNA or protein only in cells where the defective gene is expressed. Another example of a therapeutic agent is for providing a necessary product only in a certain type of cell (for example cells originating from white blood cells stem line which do not produce a specific enzyme), and in such a case the co-factor is any type of protein which characterizes these cells.

The present invention further provides an insertion vector for inserting at least one nucleic acid construct into a cell the construct comprising a heterologous coding nucleic acid sequence and an allosteric oligonucleotide which allosteric oligonucleotide has essentially no catalytic activity in the construct but can complex with a co-factor to become catalytically active; catalytic activity of the allosteric oligonucleotide regulating the expression of the heterologous coding nucleic cid sequence; wherein said co-factor is a molecule or a moiety within a molecule found inside cells.

Preferably the allosteric oligonucleotide comprises an allosteric oligonucleotide which is *a priori* inactive but which becomes active in the presence of a specific co-factor protonucleozyme being a nucleic acid molecule or complex of nucleic acid molecules having essentially no catalytic activity but can complex with a co-factor to form a nucleozyme which possess catalytic activity, the

protonucleozyme lacking a component essential for the catalytic activity of the nucleozyme and said co-factor provides said component.

BRIEF DESCRIPTION OF THE DRAWINGS

In order to understand the invention and to see how it may be carried out in practice, a preferred embodiment will now be described, by way of non-limiting example only, with reference to the accompanying drawings, in which:

Fig. 1 is a schematic representation of an array of pre-catalytic oligonucleotides on a solid support used for detection;

Fig. 2 is a schematic representation of an array of pre-catalytic oligonucleotides present in an array of liquid-holding containers used for detection;

Figs. 3A, 3B and **3C** are a schematic representation of an array of candidates for catalytic oligonucleotides present in an array of containers used for *in vitro* evolution, wherein each step of the method selected conditions become more stringent;

Fig. 4 is a schematic representation of a plurality of arrays of candidates for catalytic oligonculeotides present on a plurality of solid supports used for *in vitro* evolution;

Figs. 5A-F show a schematic representation of the method for determining or monitoring *in vivo* or *in situ* analytes in cells; and

Fig. 6A shows a method for regulating heterologous expression by allosteric nucleic acid sequences for detection purposes; and **Fig. 6B** shows a method as a GA used for therapeutic purposes;

Fig. 7 shows a gel analysis proving splicing of a protonucleozyme in the presence of a target.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

Fig. 1 shows an array of the invention **1** present on a solid support **2**, which is a glass slide covered with nucleic acid holding gel. On said gel, there are deposited in a predefined manner, thirty locations of 30 different species of

pre-catalytic oligonucleotides **3**, arranged in five columns (1 to 5) and eight rows A to F.

In each of locations **3**, there is deposited, a specific and known sequence of pre-catalytic oligonucleotides. These pre-catalytic oligonucleotides are of the type which have a missing region, missing from their core or stem, that is completed by an epitope of a specific protein, or by a specific nucleic acid sequence which is to be assayed as defined in WO 98/08974. All five locations of each row, for example locations 1, 2, 3, 4 and 5 of row A are deposited with five pre-catalytic species which become catalytically active in the presence of five different epitopes of the same protein or nucleic acid sequence. Thus, only if all of five locations 1 to 5 of a specific row are catalytically active, this indicates that the specific protein or nucleic acid sequence is present in the sample.

In this specific example, the array was incubated with a sample, which is to be assayed simultaneously for the presence of six different proteins A to F. The pre-catalytic oligonucleotides were *a priori* catalytically inactive. Each oligonucleotide contains a substrate for auto-catalytic activity of ligation of a fluorescent being labeled thereto. Fluorescent label is schematically indicated in the figure by a dark shade. The array was incubated with the sample under conditions enabling catalytic activity, and incubation takes place in the presence of a label, which is a fluorescent-bearing short sequence of oligonucleotides which can be ligated only to oligonucleotides which have become catalytic, i.e. only with those oligonucleotides which their co-factor was present in the sample. Then the array was thoroughly rinsed to remove unattached fluorescent label.

As can be seen, in the sample were present proteins C and E, as all five epitopes 1 to 5 of each of these proteins showed a fluorescent label. In addition, there was present a protein, which shared some epitopes (epitopes 2 and 4) with protein A. However, since not all of the locations of the five epitopes of protein A were labeled, this means that protein A itself was not present in the sample, but only proteins having some shared epitopes. The results may be read manually, or alternatively may be read by a CCD camera, which is connected (via an analog-to

digital converter) to a computer which can give quickly a *"yes/no"* answer only if all five epitopes of the protein were present.

Fig. 2 shows an array of containers **20**. The array is a multi-well device **21**, containing thirty wells **22**, each well holding a liquid in which there is present one species of pre-catalytic oligonucleotides **23**. The wells are arranged in columns 1 to 5 and rows A to F, as explained for Fig. 1, and each well, in fact holds exactly the same oligonucleotide species as the corresponding location in the solid support of Fig. 1. The analysis of the results is the same as described above in Fig. 1.

Fig. 3 shows a schematic representation of the *in vitro* evolution aspect of the invention.

In step (I) of the method (Fig. 3A), an array of fifteen containers is prepared, each one holding one species of synthetically produced oligonucleotides, which serve as candidates for a catalytic oligonucleotide. The oligonucleotides all show a common conserved sequence, derived from a state-of-the-art ribozymes and have also a random or semi-random variable sequence (semi-random means a sequence which showed a certain similarity, for example 70% similarity with a known sequence and is usually prepared by a technique termed *"dopping"*). It is, for example, desired to select for oligonucleotides which have catalytic activity at a temperature of 4°C. The candidate for catalytic oligonucleotides are placed in room temperature at 20°C, and it is determined in which containers catalytic activity, for example ligation with a fluorescent label is evident, as for example, can be identified by fluoresence labeling (represented schematically by a dark shade). In step (I), these containers are A1, A3, A4, B1, B2, B3, C2, C3 and C5.

Using a computer controlled automatic pipeter the liquid in these nine containers is picked up. The oligonucleotides undergo a reconstruction step, where the sequence which was ligated is not produced again, in order to obtain oligonucleotides capable of undergoing a further cycle of ligation.

Then a new array is arranged as shown in step (II) (Fig. 3B). This nine container new array, is then subjected to a temperature of 10°C and it is determined which oligonucleotides in which containers have catalytic activity in this temperature. In the present example, it is container a(I), b(II), b(III) and c(I) and c(II). The liquid of these containers is again picked up by a pipeter, the sequences reconstructed (removal of the ligated sequence) and transferred to a new array of containers having five containers. The new array in step (III) (Fig. 3C) is then subjected to a temperature of 4°C and it is clearly shown that only containers **20** and **40** hold the desired oligonucleotides having catalytic activity at this temperature. Then, the sequence of oligonucleotides **20** and **40** may be determined, by a nucleic acid sequence and they may be mass produced.

The *in vitro* evolution method is represented, for convenience means as carried out by use of an array of containers. However in practice since usually for true *in vitro* evolution purposes, where it is desired to select from many millions of different species, an array of oligonucleotides present on a solid support is the practical choice.

In such a case, it is not necessary to pick up physically the oligonucleotides in order to create new arrays (II) and (III). For example, the computer stores in its memory, the sequence of each oligonucleotide present in each one of the fifteen locations A1-A5, B1-B5, C1-C5 in the original array (I). Then, after the computer determines that only the specific nine locations A1, A3, A4, B1, B2, B3, C2, C3 and C5 show catalytic activity, a nucleic acid synthesizer can automatically creates a new array (II), having only these nine species of oligonucleotides. When the new array is subjected to another selecting condition, again the computer can read in which locations catalytic activity took place, and since it stores in its memory what is the sequence of each location, the nucleic acid synthesizer it can again prepare only these oligonucleotides in order to prepare the third array (III). The oligonucleotides may be created by the computer separately, i.e. not on the chip and then deposited on the chip using a computer-controlled robot. Alternatively, the oligonucleotides may be

synthesized in a computerized-controlled manner, *in situ* on the solid support itself. By another alternative, the sequences in each location may be obtained from a single clone prepared by genetic engineering methods. The computer has to store from which clone the sequences of each location were obtained, rather than the sequence themselves. Later, after several selective cycles, the clone may be used for mass production of the sequence.

Reference is now made to Fig. 4, which shows another embodiment in accordance with the *in vitro* evolution aspect of the present invention.

In accordance with this embodiment of the invention, a plurality of identical arrays of oligonucleotide, candidates for catalytic activity, are prepared, for example, in the present drawing 20 such identical arrays. The arrays hold oligonucleotides which are candidates for catalytic activity (which is ligation with a fluorescent bearing sequence) immobilized in a gel on a glass plate.

Each array holds fifteen different locations, bearing fifteen different species of candidate oligonucleotides. All oligonucleotides share a common conserved sequence, as well as a random or semi-random sequence which is to be selected. In practice, each array is in fact a biochip which holds 10,000-1,000,000 different species, and a plurality of arrays, for example a 100 or 1,000 which numbers are all identical if prepared.

Half of the arrays are exposed to a desired selection consisting for example the presence of the protein bovine serum albumin BSA (i.e. a catalytic activity should be evident in the presence of this molecule which serves as a co-factor). The other half of the arrays, is exposed to lack of the desired condition, i.e. catalytic activity should not be present in the absence of BSA.

For selection purposes, a *"cut-off percentage"* criteria is set both for the positive selection step (present BSA) and negative selection (absence of BSA).

In the present case, for example, the positive criteria, is that at least 80% of the locations are active in the presence of BSA, and not more than 10% of the locations are active in the absence of BSA.

The results are summarized in Table 1 shown below.

Table 1

Position	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	C1	C2	C3	C4	C5
(+) BSA	90%				60%							90%			100%
(-) BSA	30%				20%							10%			60%

As can be seen, in the present case, only position C2 fulfills the above cut-off criteria, as more than 80% of the array should be active in the presence of BSA, and no more than10% of the arrays should be active in the absence of BSA. In usual *in vitro* evolution methods, which are carried in a test tube, the selection criteria are gradually made more stringent with each round of selection. For example, where the desired activity is the presence of a certain compound such as BSA, then in each round of selection the concentration of the BSA is lowered. This is done since it is very difficult to discriminate in the beginning between *"noise"* and *"signal"* so that in order not to lose the signal, the criteria is initially quite liberal, and only when some noise is eliminated, then the criteria may be made more stringent.

Against this, when using an *"array-based in vitro evolution selection"* of the invention since there is essentially very little background noise, and each location is determined separately, it is possible to start the selection with the most stringent criteria desired. For example, in the present case, it is possible to expose the arrays to a very low concentration of BSA. If there is not enough signal, or not enough locations to show up a catalytic activity, then it is possible to gradually raise the concentration of BSA until a satisfactory number of locations are catalytically active. Thus, in fact theoretically it is possible to select for catalytically active oligonucleotides, from a huge number of possible candidates, using a single step.

It is also possible to carry out selection for oligonucleotides which show catalytic activity in the presence of two selective conditions, either by exposing the arrays to the two conditions at once, or by using for example four or five duplicate sets of arrays, and then selecting locations which are active above a *"cut-off percentage"* in all the selected conditions and are active in less than a cut-off percentage in all non-selected conditions. For example in the present case, it is preferable to try and carry out negative selection also in the presence of a protein which is not BSA.

Fig. 5, shows the method for determining, or monitoring *in situ* or *in vivo* analytes inside cells.

Fig. 5A shows the pre-catalytic oligonucleotide **1.** The oligonucleotide has an analyte binding region **2**, which can bind the analyte to be detected, for example, a certain protein produced by viruses which indicates the cell is infected by said virus. The oligonucleotide **1** has a sequence **3**, which serves as a substrate for cleavage. In addition, the oligonucleotide carries two moieties which enable detection, the first moiety **4** is a fluorescent label such as fluorescein. The second, is a wavelength modulator **5**. When modulator **5**, and fluorescent label **4** are adjacent to each other, such as shown in Fig. 5A, the wavelength fluorescence of **4** is shifted, and if the fluorescent rendered is set to the unshifted wavelength no signal is detected. In addition, the oligonucleotide carries identification *tag* sequence **6**.

A construct for inserting the oligonucleotides **1** into cell **7** (Fig. 5B) is prepared. The construct may be a plasmid, a carrier viral vector, etc.

If the analyte **8,** for example a viral protein, is present in the cell, it binds to analyte-binding site **2** (Fig. 5C). Then the pre-catalytic oligonucleotide converts to its catalytically active form due to binding of analyte **8** (Fig. 5D). In its active form, it performs auto-cleavage, releasing to the medium molecule **9** comprising the fluorescent label **4** and identification *tag* sequence **6** (Fig. 5E)

In the case where only a single analyte is to be detected, the fact that the fluorescent label **4** is distanced from modulator **5**, produces a shift in the fluorencent wavelength emitted so that the wavelength may be read and detected in the cell, or in the entire tissue, since fluorescence can be emitted through living membranes and hence read also in whole tissues. Such a method enables real-time monitoring in live tissue of analytes. In a case where a single analyte is to be detected, of course the oligonucleotide **1** does not have to carry an identification *tag* sequence, which is to be used only where a plurality of analytes are to be detected simultaneously. In such a case the construct for inserting the oligonucleotides into the cells (plasmid, viral vector) carries a plurality of

different oliognucletoides **1**, each having a slightly different analyte-binding region **2**, which can be activated by different analytes.

Fig. 5F, shows an embodiment of the detection or determination assay used to detect a plurality of different analytes simultaneously. In such a case the construct for inserting the oligonucleotides into the cells (being a plasmid, viral vector, etc.) carries a plurality of different oligonucleotides each one having a slightly different analyte-binding region **2** which can be activated by different analytes

The figure shows a chip **10** having schematically three short sequences thereon **11, 12** and **13** each complementary to a different identification tag sequence of the original pre-catalytic oligonucleotides **1**.

Molecules **14** and **15** are short sequences cleaved from oligonucleotides, (corresponding to molecule **9** in Fig. 5E) bearing fluorescent labels which were cleaved due to the fact that the oligonucleotides became catalytically active in the presence of a suitable analyte which served as their co-factor. Oligonucleotide **16**, which bind to complementary sequence **13** has a wavelength due to the modulator moiety being adjacent to the fluorescent label so that if the fluorescent reader is set to a specific wavelength no fluorescent is read.

When reading the chip, one can see that fluorescence is present in locations **11** and **12**, while no fluorescence is read (due to a shift in wavelength) in location **13**. By knowing exactly which sequences are present in locations **11, 12** and **13**, it is possible to know which complementary identification tag sequences are present on molecule **14, 15** and on oligonucleotide **16**, and thus it is possible to know that the analyte which bound to pre-catalytic oligonucleotides originally attached to molecules **14** and **15** was present in the cell, while the analyte which should have activated pre-catalytic oligonucleotide **16** is not present in the cell.

Reference is now made to Fig. 6A which shows the aspect of regulating the expression of a heterologous gene, for detection purposes. A nucleic acid construct (top representation) is built comprising a gene coding for a detectable

produce for example a gene coding for the luciferase enzyme is broken into two parts (shown schematically as ½R) with the allosteric oligonucleotide intervening thereinbetween. The allostric oligonucleotide in the present case is of the type of protonucleozyme as specified in WO 98/08974 which has a missing component which can be completed by a suitable co-factor. The construct is inserted (for example by a vector) into cells (second representation. The insertion is indiscriminate, i.e. all cells may contain this construct). The co-factor is, for example, a viral protein present only in cells infected with the virus. Thus only in these cells, is the missing component of the protonucleozyme completed by the co-factor (third representation), while in other cells due to lack of correct co-factor the probe nucleozyme is not completed. The completed protonucleozyme becomes active, thus splicing itself out from the heterologous nucleic acid sequence (fourth representation) enabling its expression to produce the luciferase enzyme – (indicated by **r**) which produces a detectable label.

Fig. 6B shows essentially the same concept, but in that case the heterologous nucleic acid sequence is one coding for a therapeutical agent (marked schematically as ½T). For example, this agent may be a toxin or a cytotoxic agent such as catapase. Only if the suitable co-factor, for example a viral protein, is present in the cells is the allosteric oligonucleotide is spliced out, and the cytotoxic gene may be expressed producing the product (**t**) leading to death only of these cells.

Fig. 7 shows the results of splicing catalytic activity obtained by a protonucleozyme capable of self-splicing only in the presence of a specific co-factor. A series of protonucleozymes derived from a variety of Group I intron ribozymes (referred to in Table 2 below) were engineered. The protonucleozymes hybridize to a specific co-factor resulting in an active ribozyme (nucleozyme). Its activity resulted on the excision followed by the ligation of its 5′ and 3′ exons, (i.e. splicing).

Radioactively labeled GTP (needed for the reaction to proceed) was used to follow the activity of the constructs. The GTP binds to the resulting intron (see Fig. 7) producing a size shift that is detected on gel electrophoresis (Fig. 7).

Table 2

Protonucleozyme	Source of the Group intron
1	Bacteriophage T4 nrdL
2	Bacteriophage T4 td
3	Bacteriophage T4 td (without its P7.1, and P7.2 region)
4	Bacteriophage T4 nrdB
5	Hybrid of Bacteriophage T4 nrdB and Bacteriophage T4 sunY
6	Bacteriophage T4 sunY
7	Anabaena sp

Experimental Method

4 μm of protonucleozyme with its specific co-factor were pre-incubated 25 mM HEPES buffer, pH-7.3; 30 mM $MgCl_2$ at 50° C for 15 min. The reaction was started by the addition of 100 μM GTP (spiked with 1:300 α-^{32}P-GTP) and incubated for 15 minutes and 60 minutes at 32°C.

Controls were carried out in the absence of co-factor for each one of the specific pronucleozyme.

The reactions were stopped by the addition of 2x loading buffer in the presence of 7M urea, heated at 80°C for 5 minutes and run on 20% AGE in the prsence of 7M urea.

As can be seen, under these conditions (i.e. presence of co-factor) the pronucleozyme was able to excise and ligate (splice).

CLAIMS:

1. An array of catalytic oligonucleotides, pre-catalytic oligonucleotides, or candidates for catalytic oligonucleotides immobilized on a solid support, wherein each oligonucleotide-bearing location of the solid support, bears essentially a single species of said oligonucleotides.

2. An array of catalytic oligonucleotides, pre-catalytic oligonucleotides, or candidates for catalytic oligonucleotides present in an array of liquid-holding containers, wherein each container holds essentially a single species of said oligonucleotides.

3. An array according to Claim 1 or 2, wherein the pre-catalytic oligonucleotides are oligonucleotides which are *a priori* inactive, and which become catalytically active in the presence of a co-factor.

4. An array according to Claim 3, wherein the pre-catalytic oligonucletoides are proto-nucleozymes, being nucleic acid molecules or a complex of nucleic acid molecules, having essentially no catalytic activity but which can complex with the co-factor to form nucleotzymes which possess a catalytic activity, the proto-nucleozymes lack a component essential for the catalytic activity of the nucleozyme and said co-factor provides said component.

5. An array according to Claim 4, wherein said missing component in the proto-nucleozyme is a missing segment of one or more nucleotides.

6. An array according to Claim 4 or 5 wherein the nucleozyme is a ribozyme.

7. An array according to Claim 1 or 2, wherein the oligonucleotides, being candidates for catalytic activity, are oligonucleotides having a conserved sequence, common to all species in the array, which is identical to a corresponding region of a catalytically active ribozyme and at least one random sequence, being variable in different species of the oligonucleotides.

8. An assembly for the detection of the presence of an analyte in a sample comprising an array according to Claim 1 or 2, and a signal producer capable of producing a detectable signal in locations of the array which hold oligonucleotides having catalytic activity.

9. An assembly according to Claim 8, wherein the detectable signal is fluorescence.

10. An assembly according to Claim 8, wherein the detectable signal is magnetic force.

11. An assembly according to Claim 8, wherein the signal producers are a nucleic acid sequence contained in the oligonucleotides, which sequence changes at least one detectable property upon auto-catalytic activity in *cis*.

12. A system for the detection of the presence of a co-factor in a sample comprising the assembly of Claim 8, and a detector capable of reading the detectable signal.

13. A system according to Claim 12, wherein the detectable signal is fluorescence, and the detector is a CCD camera.

14. A system according to Claim 13, further comprising an analog-to-digital convector capable of transducing the CCD output to digital information, and a computer capable of processing said information.

15. A method for detecting the presence of at least one analyte in a sample, each analyte being the co-factor of one specific species of pre-catalytic oligonucleotides as defined in Claim 5, the method comprising

(i) providing an array according to Claim 1, comprising pre-catalytic oligonucleotides;

(ii) incubating the array with a sample under conditions allowing catalytic activity of oligonucleotides;

(iii) determining by utilizing the system of Claim 12, in which locations catalytic activity took place;

said determination indicating the specific oligonucleotide species which became catalytically active, thereby indicating the presence in the sample of the specific analytes which served as co-factors of said specific species which become active.

16. A method for detecting the presence of at least one analyte in a sample, each analyte being the substrate of a specific *trans* active catalytic oligonucleotide the method comprising:

(i) providing an array according to Claim 1, comprising catalytic oligonucleotides;

(ii) incubating the array with a sample under conditions allowing catalytic activity of oligonucleotides;

(iv) determining, utilizing the system of Claim 12, which locations catalytic activity took place;

said determination indicating the specific oligonucleotide species which became catalytically active, thereby indicating the presence in the sample of the specific analyte which served as the specific substrate of said specific catalytic species which become catalytically active.

17. A method according to Claim 15 or 16, for the simultaneous detection of a plurality of analytes in a sample.

18. A method according to Claim 15, wherein the co-factor is a non-nucleic acid molecule.

19. A method according to Claim 15, wherein the co-factor is a protein.

20. A method for selecting for oligonucleotides having a catalytic activity under a desired condition comprising:

(i) providing an array of candidates for catalytic oligonucleotides according to Claim 1;

(ii) exposing the array to the desired condition;

(iii) determining in which locations the catalytic activity took place, said location indicating the location of the oligonucleotide to be selected; and

(iv) determining the sequence of the oligonucleotides in said location, wherein said oligonucleotides have a desired catalytic activity under said conditions.

21. A method according to Claim 20, wherein the sequence of the oligonucleotides selected in step (iii) are determined in step (iv) by utilizing a computer stored database comprising the sequence of oligonucleotides present in each location.

22. A method for selecting oligonucleotides having catalytic activity under a desired condition comprising:

(i) providing a plurality of identical arrays of candidates for catalytic oligonucleotides according to Claim 1;

(ii) exposing the plurality of arrays to a desired condition;

(iii) determining in which locations catalytic activity took place in at least a pre-determining percentage of individual arrays; and

(iv) determining the sequence of the oligonucleotides in said locations.

23. A method for selecting oligonucleotides having catalytic activity under a desired condition and which do not have catalytic activity under a non-desired condition comprising:

(i) providing a plurality of identical arrays of candidates for catalytic oligonucleotides according to Claim 1;

(ii) exposing a first portion of the arrays to the desired condition and exposing the remaining second portion of the arrays to the non-desired condition;

(iii) determining in which locations catalytic activity took place under the desired condition in at least a pre-determined percentage of arrays in said first portion; and in which locations catalytic activity took place, under a non-desired activity in no more than a predetermined percentage of the array in said second portion; said locations indicating the sequence of the oligonucleotides to be selected; and

(iv) determining the sequence of the oligonucleotide in said locations.

24. An insertion vector for inserting at least one nucleic acid construct into a cell, the construct comprising an allosteric oligonucleotide that is *a priori* active, but which become active in the presence of a specific co-factor; the nucleic acid construct further comprising a detectable moiety which changes at least one detectable property upon *cis* catalytic activity of the allosteric oligonucleotide; the co-factor for the allosteric oligonucleotide being a molecule or a moiety within a molecule found in cells.

25. A vector according to Claim 24, wherein the allosteric oligonucleotide is a protonucleozyme which is *a priori* inactive but which becomes active in the presence of a specific co-factor, the protonucleozyme being a nucleic acid molecule or complex of nucleic acid molecules having essentially no catalytic activity but can complex with a co-factor to form a nucleozyme which possess catalytic activity, the protonucleozyme lacks a component essential for the catalytic activity of the nucleozyme and said co-factor provides said component.

26. A vector according to Claim 24, wherein the catalytic activity is splicing or cleavage.

27. A vector according to Claim 24, wherein the label is a fluorescence emitting moiety and a quencher.

28. A vector according to Claim 24, comprising a plurality of nucleic acid constructs, each construct comprising an allosteric oligonucleotide which becomes catalytically active through complexing with a different co-factor, each nucleic acid construct further comprising a different identification *tag* sequence unique to said construct, said identification *tag* sequence being adjacent to the detectable label and being detached together with said label, from the construct upon catalytic activity *in cis*.

29. A method for determining the presence of an analyte *in vivo* or *in situ*, in a cell comprising:

(i) utilizing the vector of Claim 24 to insert the nucleic acid molecule construct to the cell, wherein the analyte to be detected is the co-factor of the protonucleozyme; and

(ii) determining the existence of the detectable labels in the cells, said existence signifying the presence of the analyte in a cell.

30. A method according to Claim 29, wherein the identification of the nature of the label-bearing identity tags is carried out by contacting the medium with an array of immobilized sequences, said immobilized sequences being complementary to the *tag* sequences.

31. An insertion vector for inserting at least one nucleic acid construct into a cell the construct comprising a heterologous nucleic acid sequence and an allosteric oligonucleotide; the allosteric oligonucleotide having essentially no catalytic activity in the construct but can complex with a co-factor to become catalytically active; catalytic activity of the allosteric oligonucleotide regulates the expression of the heterologous nucleic acid sequence; wherein said co-factor is a molecule or a moiety within a molecule found inside cells.

32. A vector according to Claim 31, wherein the allosteric oligonucleotide which is *a priori* inactive but which becomes active in the presence of a specific co-factor, is a protonucleozyme being a nucleic acid molecule or complex of nucleic acid molecules having essentially no catalytic activity but which can complex with a co-factor to form a nucleozyme which possess catalytic activity, the protonucleozyme lacks a component essential for the catalytic activity of the nucleozyme and said co-factor provides said component

33. An insertion vector according to Claim 31, where the regulation is up-regulation and wherein the heterologous nucleic acid sequence is *a priori* constructed so as not to express a desired product; the catalytic activity of the allosteric oligonucleotide on said heterologous sequence, causing it to express a desired product.

34. An insertion vector according to Claim 31, where regulation is down-regulated and wherein the heterologous nucleic acid sequence is *a priori* constructed so as to express a desired product; the catalytic activity of the allosteric oligonucleotide on said heterologous sequence causing it not to express the desired product.

35. An insertion vector according to Claim 33, wherein the desired product is a detectable label.

36. An insertion vector according to Claim 33 or 34, wherein the desired product is a therapeutical agent.

37. An insertion vector according to Claim 36, wherein the desired product is a cytotoxic agent.

38. An insertion vector according to any one of Claims 31 to 37, wherein the co-factor is selected from the group consisting of: a viral protein, a viral mRNA, a viral DNA, a mutated cellular protein, and a mutated cellular mRNA.

39. A method for regulating the expression of a heterologous nucleic acid sequence in a desired cell population comprising:

(i) utilizing the vector of Claim 31 to insert the heterologous nucleic acid sequence to cells, wherein the co-factor is essentially specific to the desired cell population; and

(ii) providing conditions enabling catalytic activity of the allosteric oligonucleotides.

40. A method according to Claim 39, for up-regulating the expression of a heterologous sequence in a desired cell population wherein the vector is the vector of Claim 33.

41. A method of down-regulating the expression of a heterologous sequence in a desired cell population wherein the vector is the vector of Claim 34.

42. A method for detecting the presence of an analyte, *in vivo* or *in situ* in a cell comprising:

(i) utilizing the vector of Claim 35, to insert the nucleic acid construct into cells wherein the analyte to be detected is the co-factor of the allosteric oligonucleotide; and;

(ii) determining the existence of the detectable labels in the cells, said existence signifying the presence of the analyte in a cell.

43. A method for genetic therapy in a desired cell population the method comprising:

(i) utilizing the vector of Claim 36 to insert the nucleic acid construct into cells wherein the co-factor is essentially specific to the desired cell population; and

(ii) providing conditions enabling activity of the allosteric oligonucleotide.

44. A method according to Claim 43, wherein the desired cell population is selected from the group consisting of viral infected cells, cancerous cells, and cell producing mutated proteins.

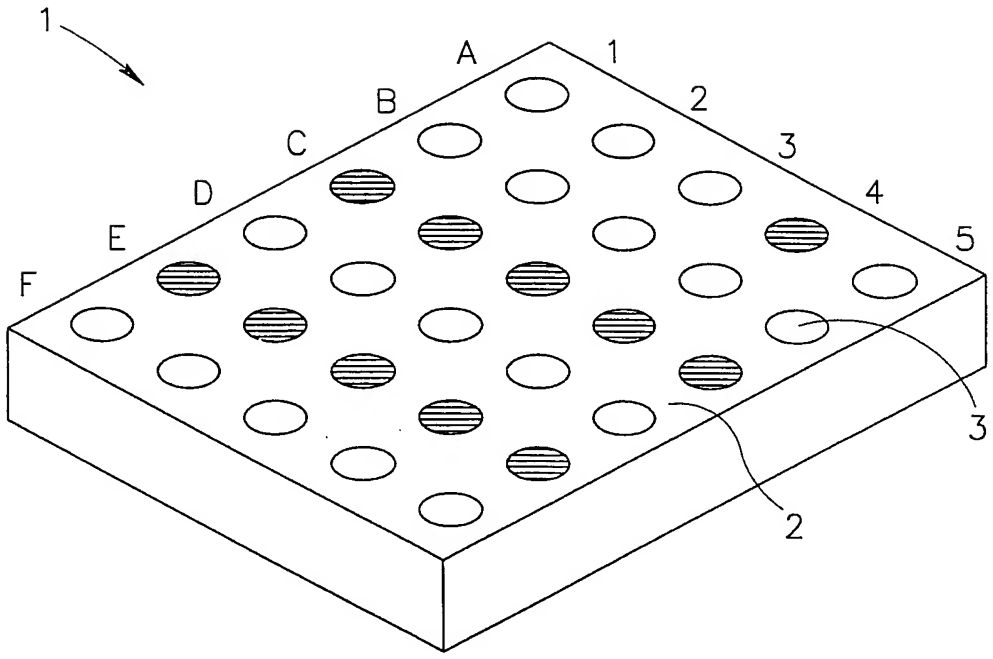


FIG.1

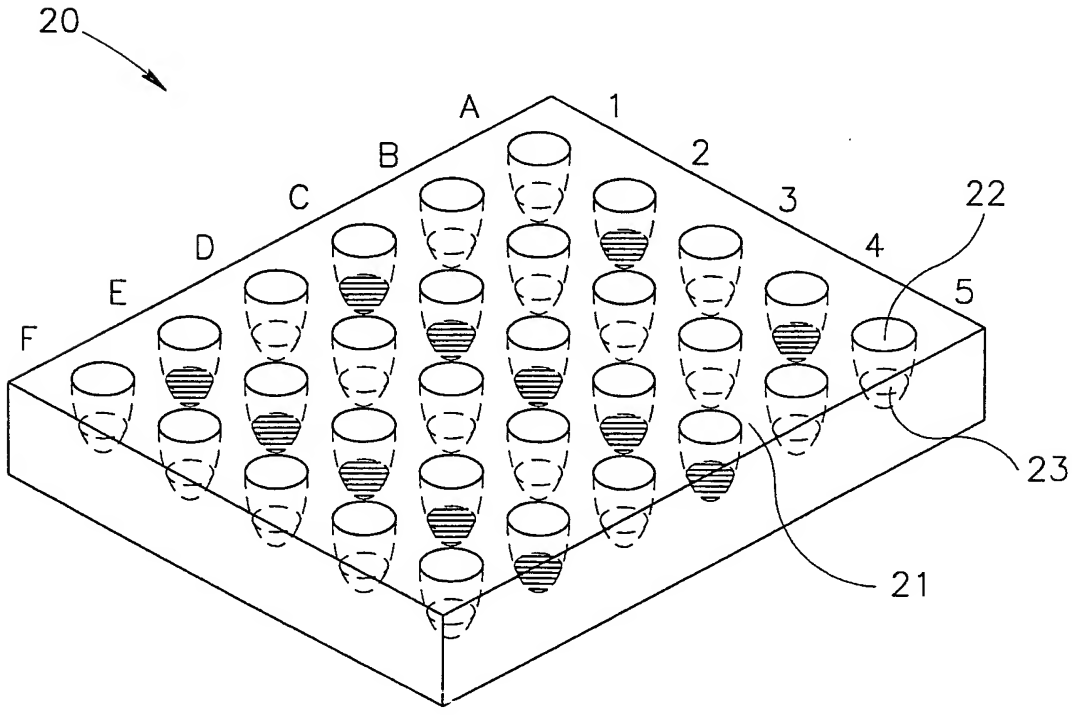


FIG.2

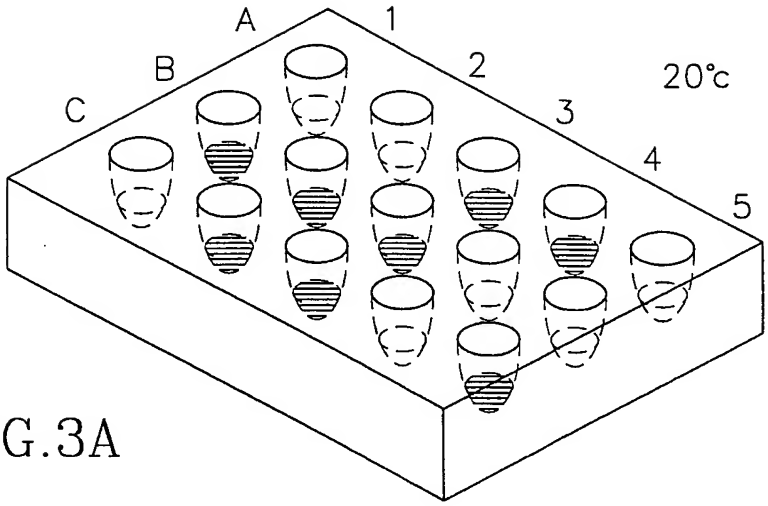


FIG.3A

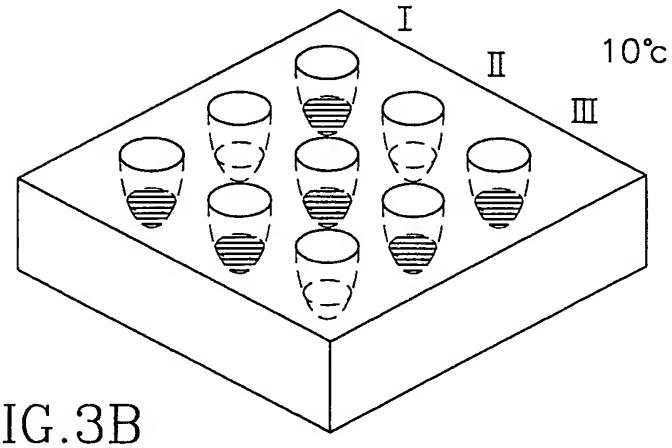


FIG.3B

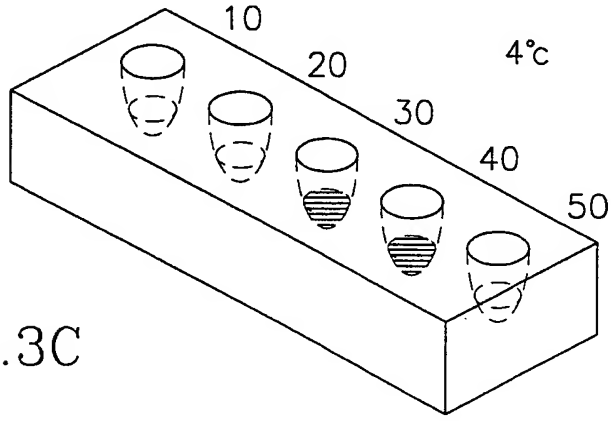


FIG.3C

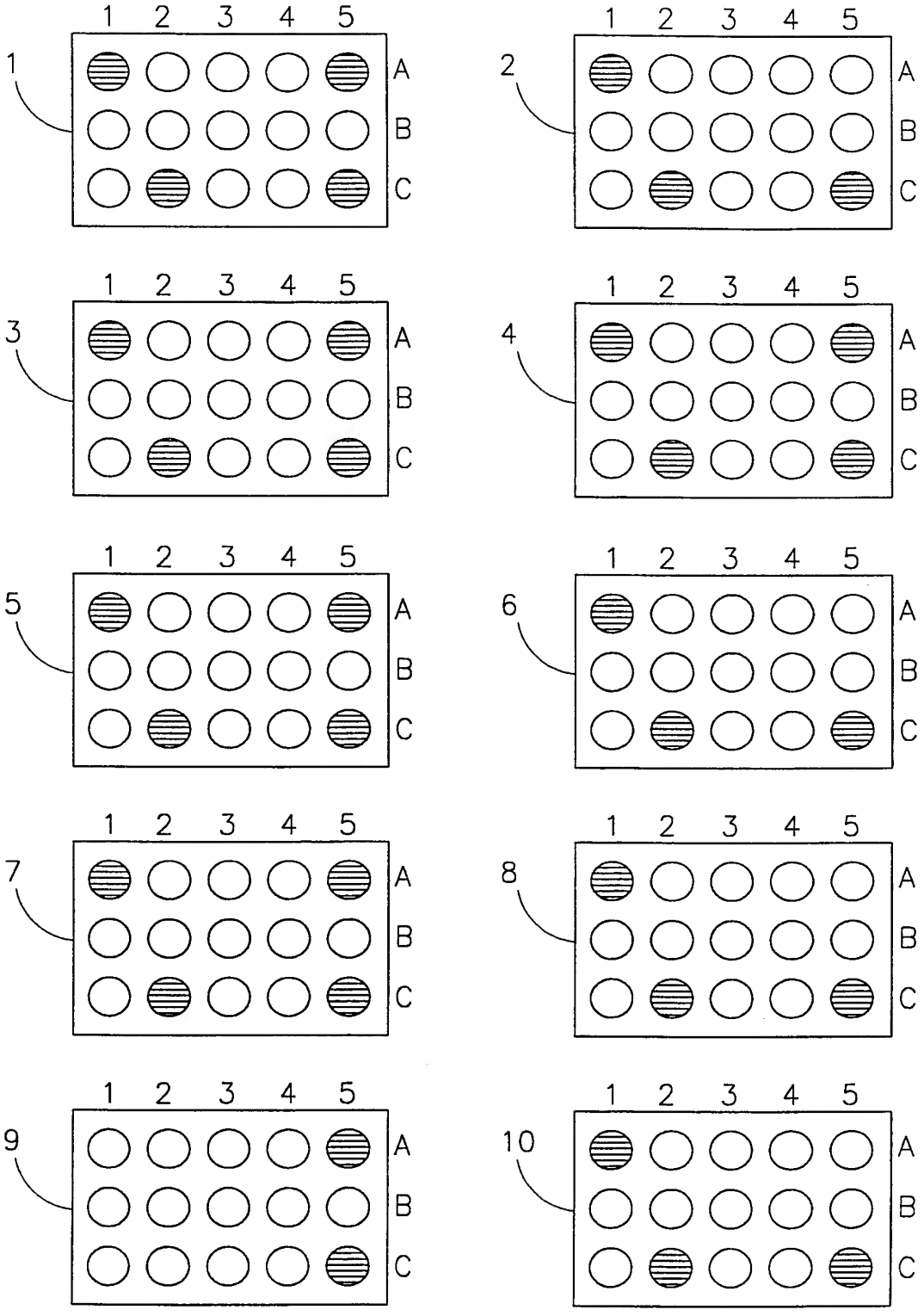

1 2 3 4 5
1
A
B
C
1 2 3 4 5
2
A
B
C
1 2 3 4 5
3
A
B
C
1 2 3 4 5
4
A
B
C
1 2 3 4 5
5
A
B
C
1 2 3 4 5
6
A
B
C
1 2 3 4 5
7
A
B
C
1 2 3 4 5
8
A
B
C
1 2 3 4 5
9
A
B
C
1 2 3 4 5
10
A
B
C

FIG.4A

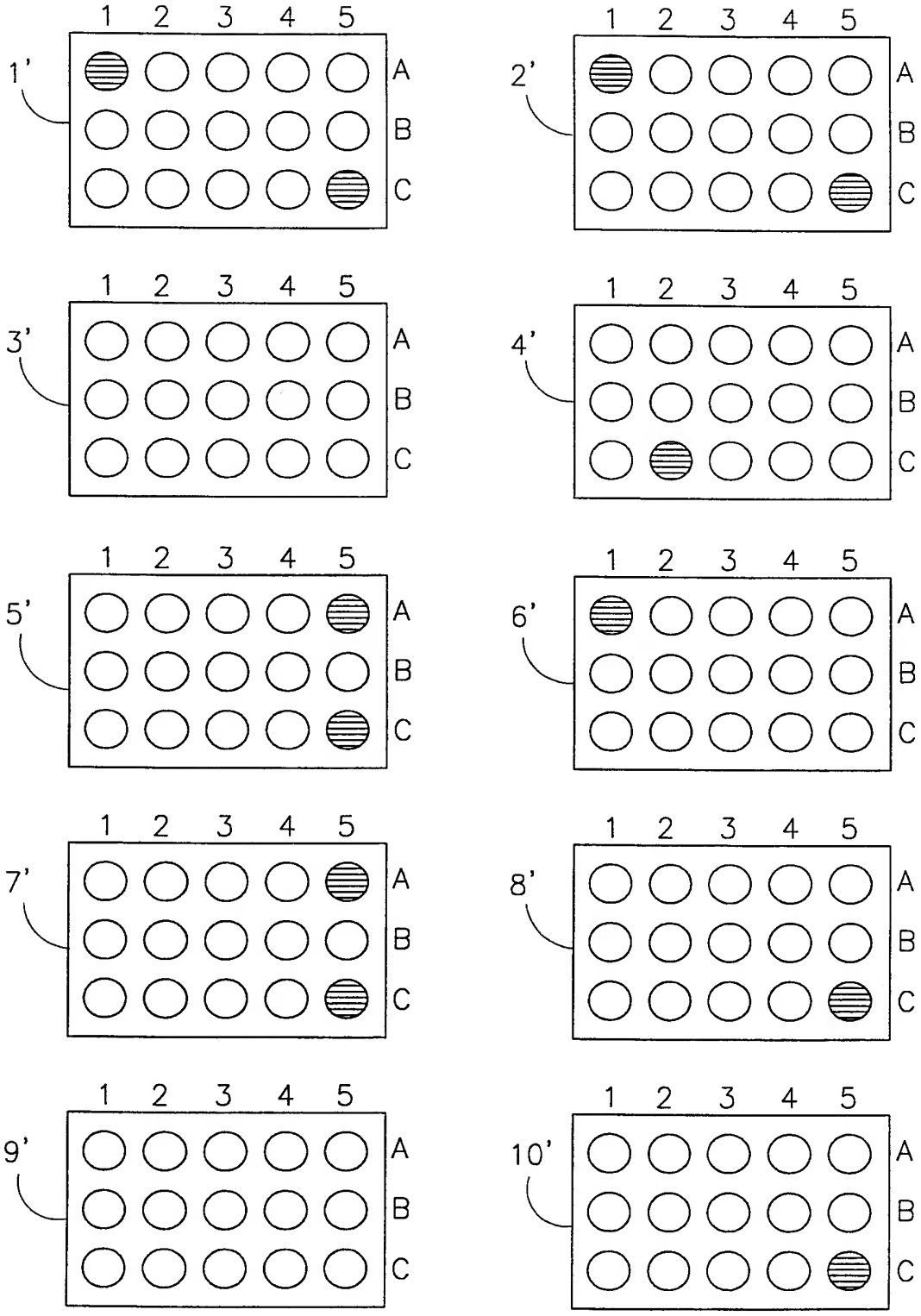

1 2 3 4 5
1'
A
B
C
1 2 3 4 5
2'
A
B
C
1 2 3 4 5
3'
A
B
C
1 2 3 4 5
4'
A
B
C
1 2 3 4 5
5'
A
B
C
1 2 3 4 5
6'
A
B
C
1 2 3 4 5
7'
A
B
C
1 2 3 4 5
8'
A
B
C
1 2 3 4 5
9'
A
B
C
1 2 3 4 5
10'
A
B
C

FIG.4B

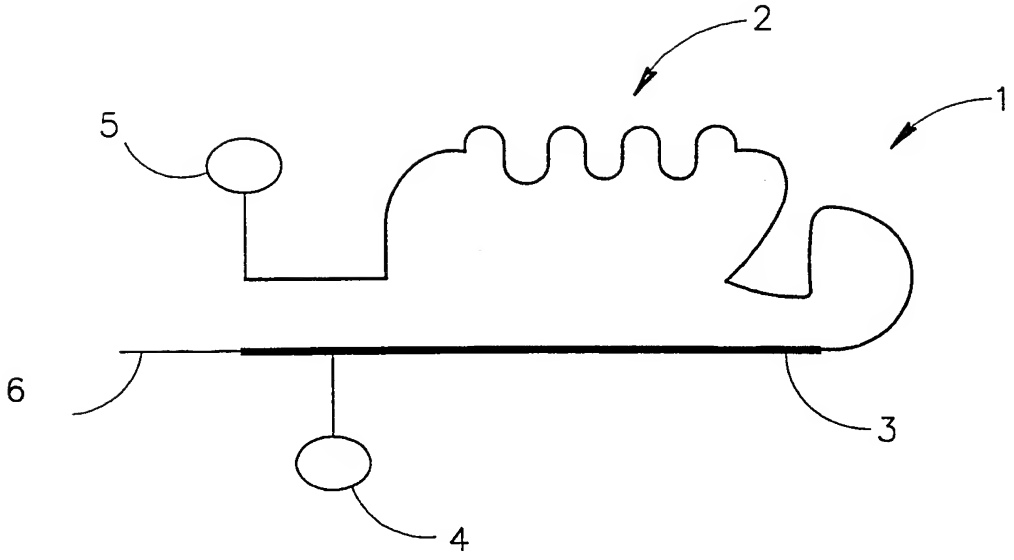


FIG.5(a)

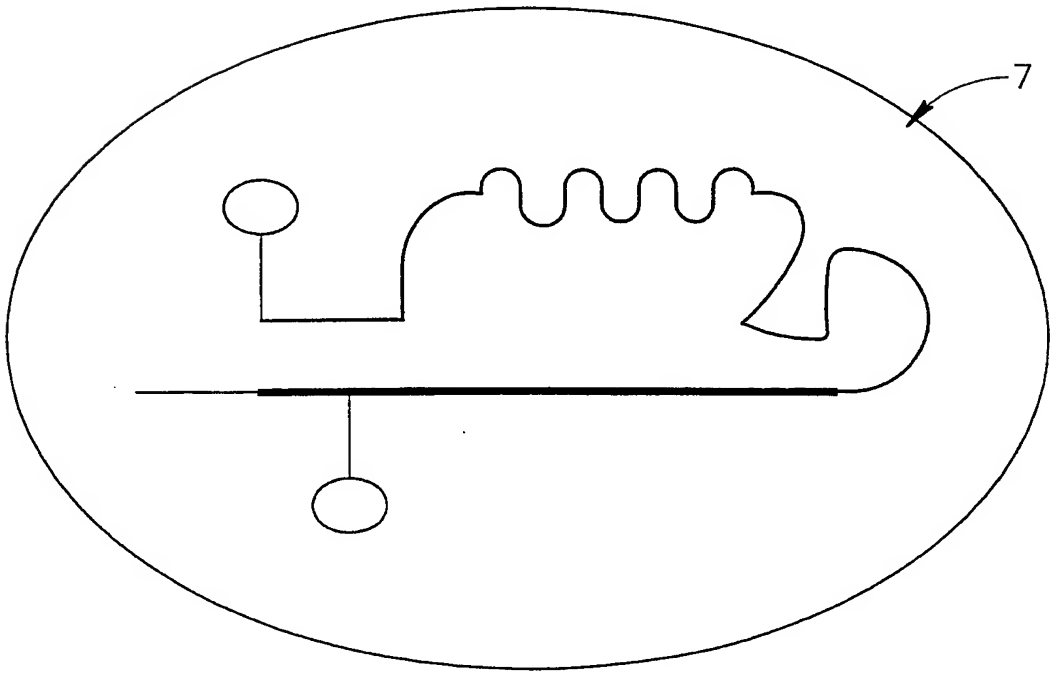


FIG.5(b)

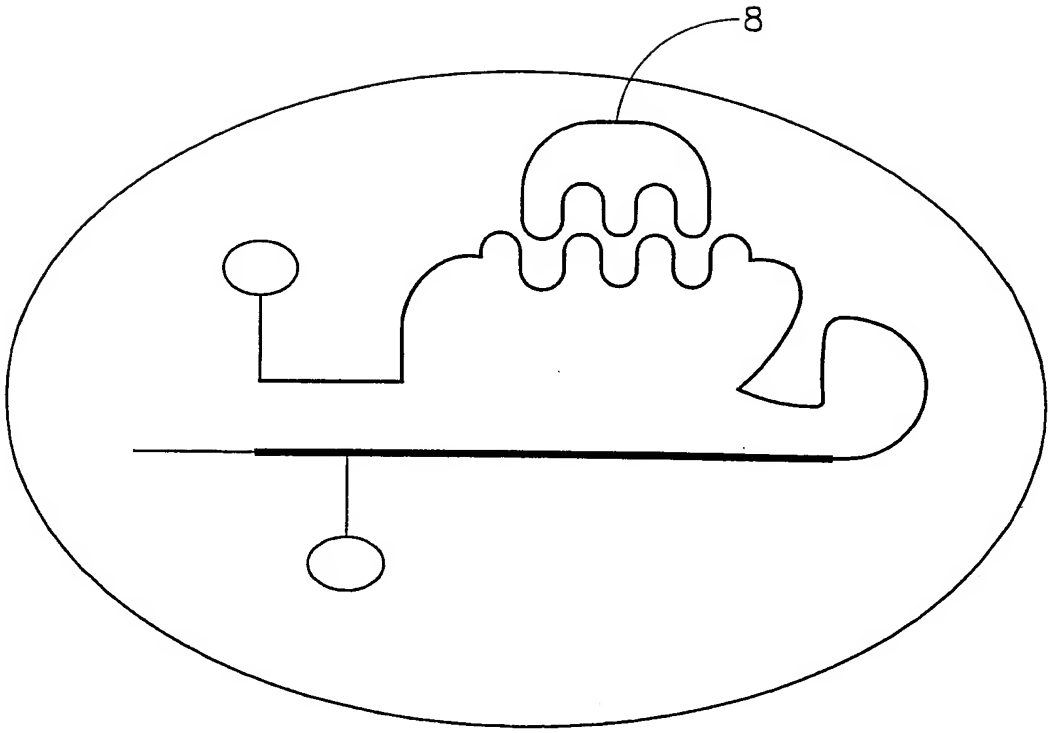


FIG.5(c)

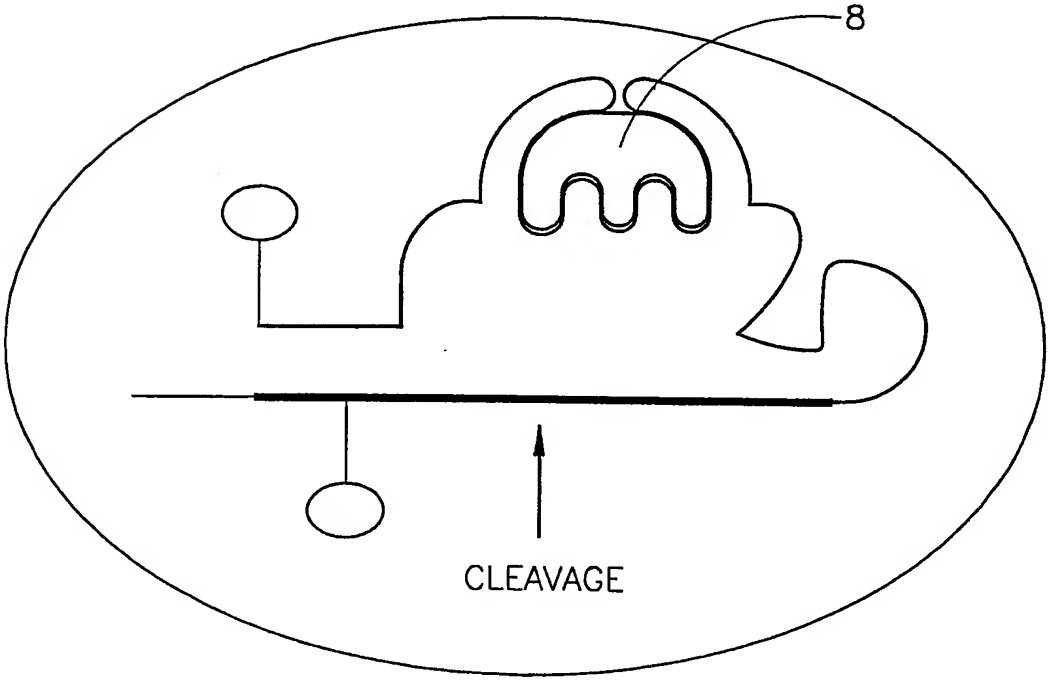


FIG.5(d)

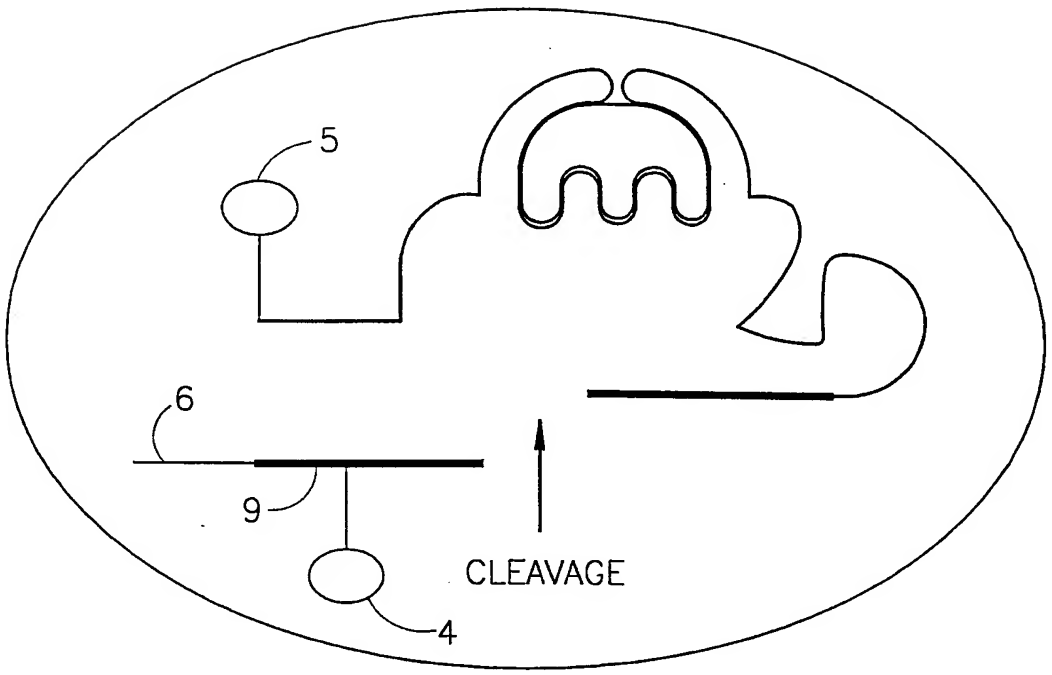

5
6
9
CLEAVAGE
4

FIG.5(e)

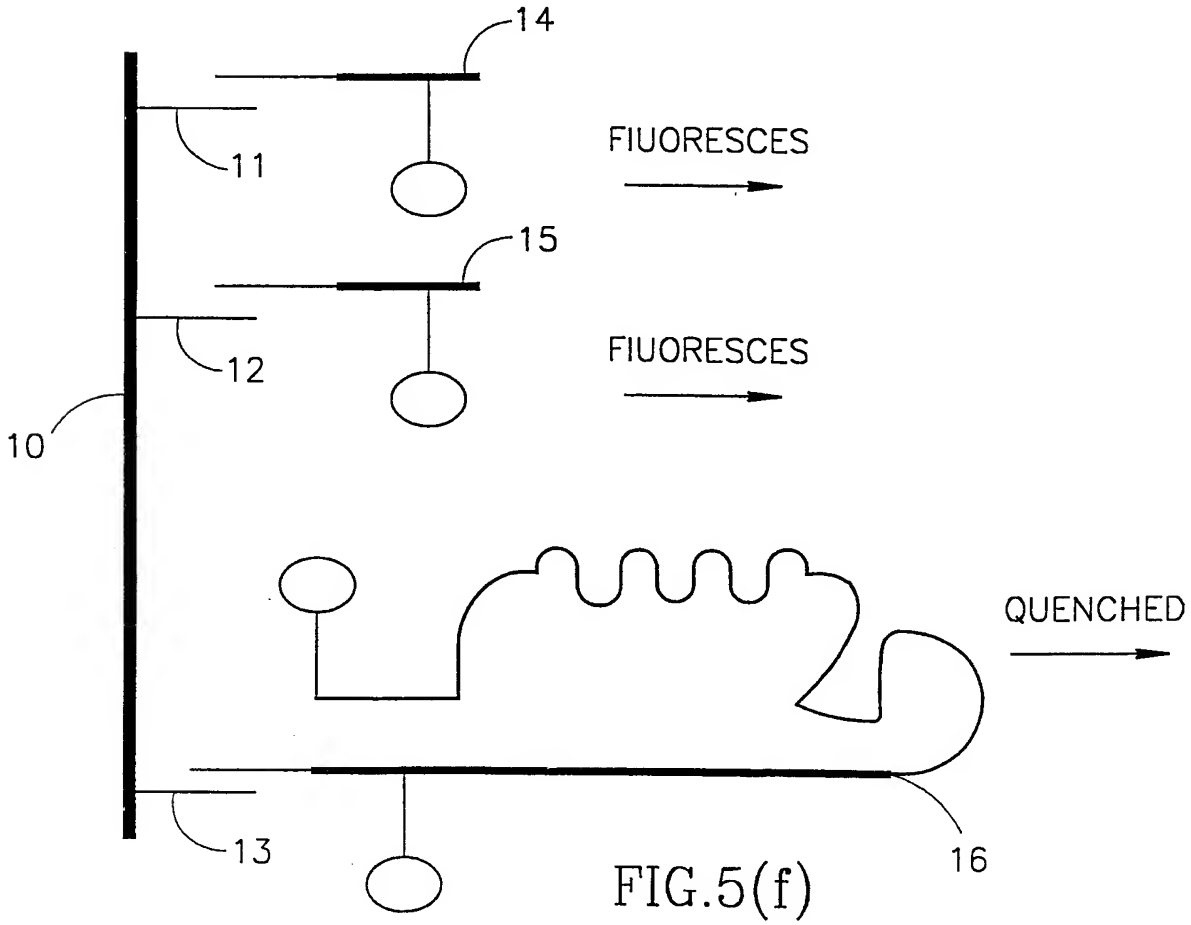

14
11
FIUORESCES
15
12
FIUORESCES
10
QUENCHED
13
16

FIG.5(f)

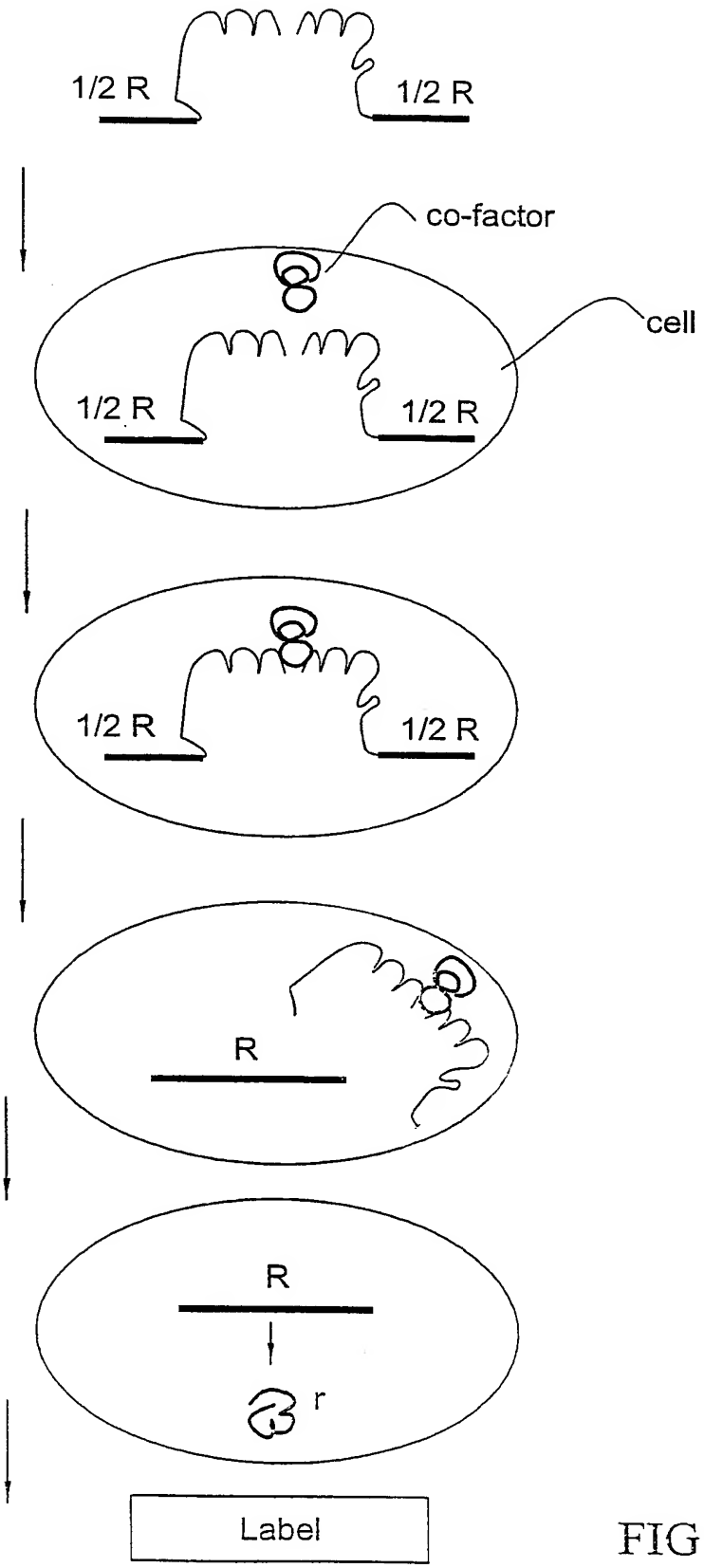


FIG. 6A

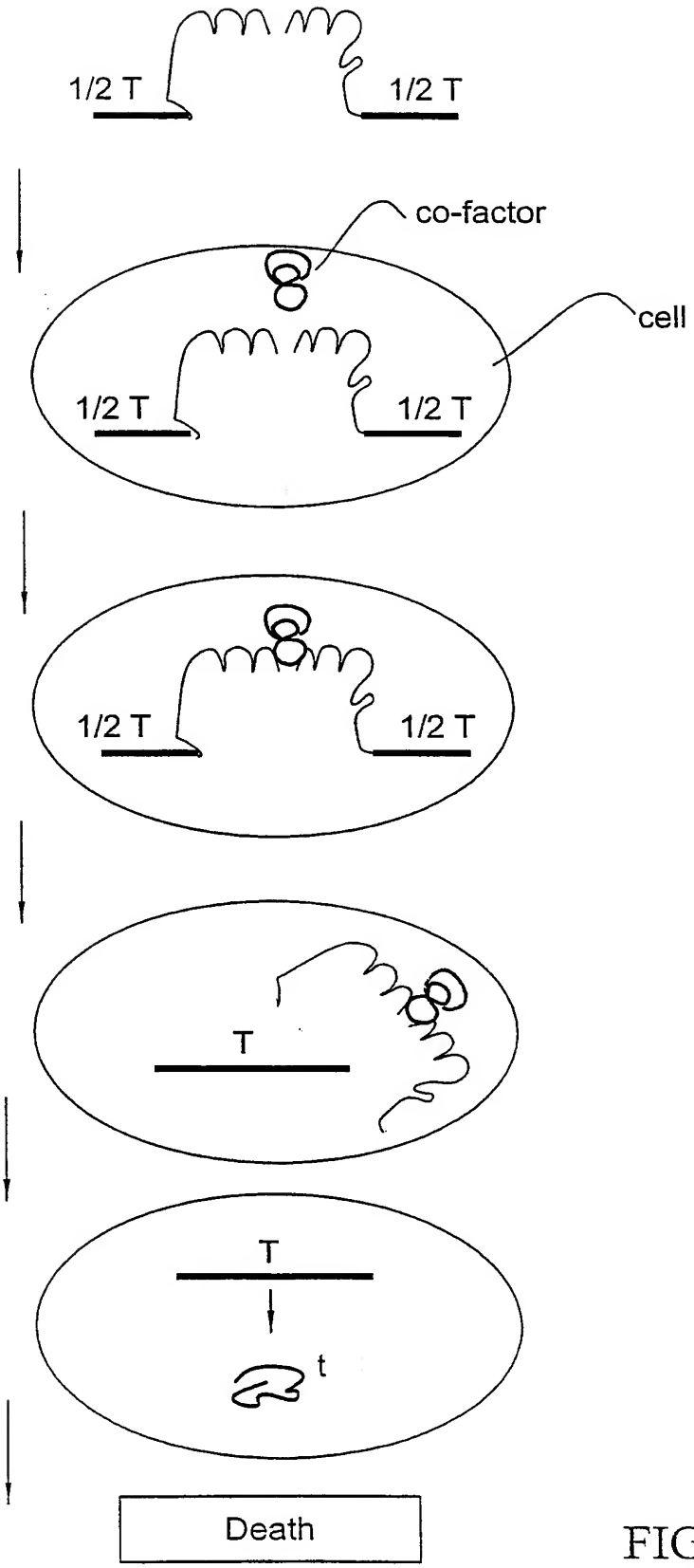


FIG. 6B

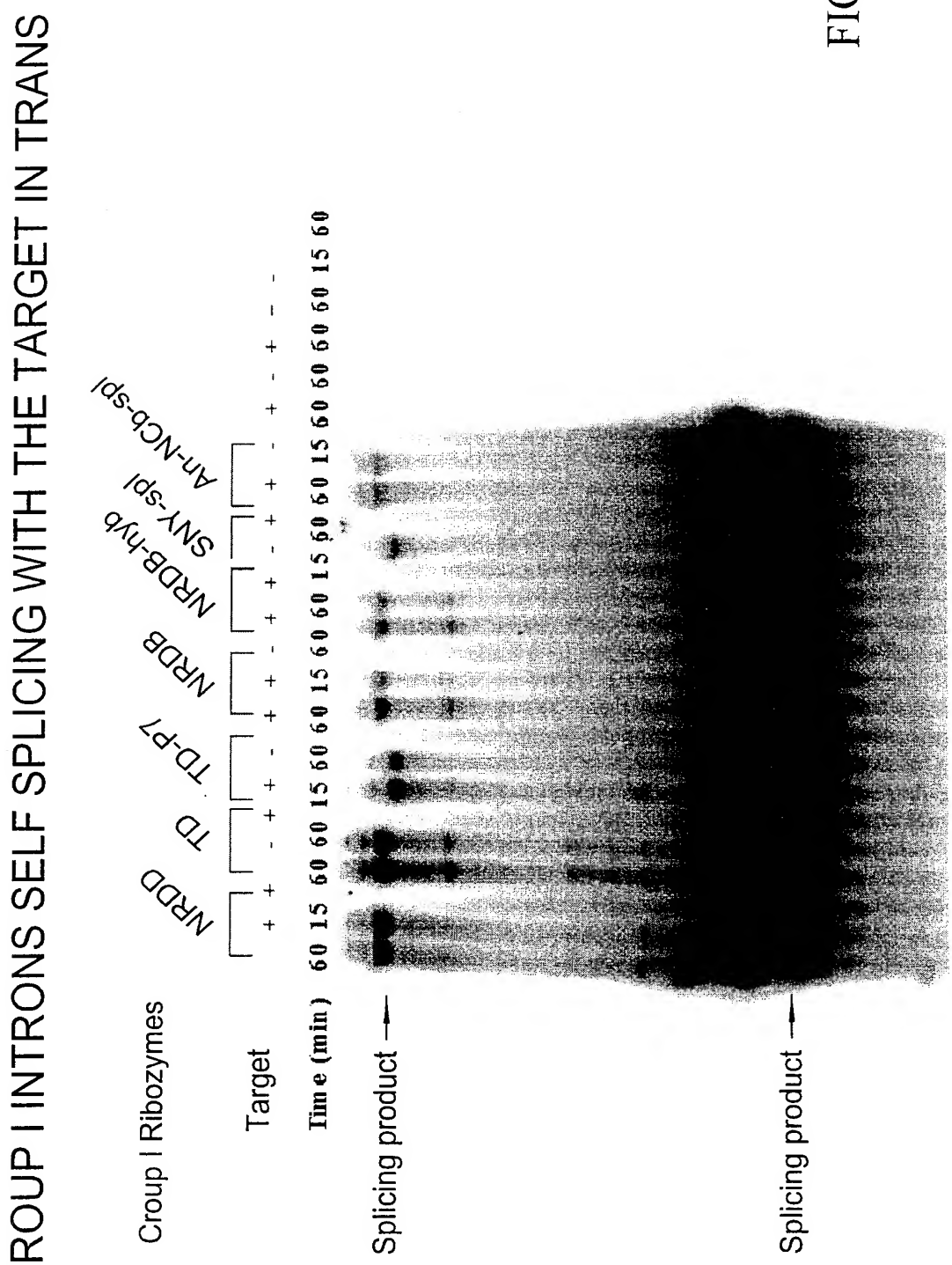

GROUP I INTRONS SELF SPLICING WITH THE TARGET IN TRANS
Croup I Ribozymes
NRDD
TD
TD-P7
NRDB
NRDB-hyb
SNY-spl
An-NCb-spl
Target
Time (min)
60 15 60 15 60 60 15 60 60 15 60 60 15 60 60 60 60 60 15 60
Splicing product
Splicing product

FIG. 7